```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:   September 16, 2005, 16:54:01 ; Search time 117 Seconds
                                          (without alignments)
                             786.743 Million cell updates/sec

Title:         US-10-757-624-3
Perfect score: 1276
Sequence:      1 MSKGEELFTGVVPILVELDG.........VLLGFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:    2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  A_Geneseq_16Dec04:*
         1: geneseqp1980s:*
         2: geneseqp1990s:*
         3: geneseqp2000s:*
         4: geneseqp2001s:*
         5: geneseqp2002s:*
         6: geneseqp2003as:*
         7: geneseqp2003bs:*
         8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1276	100.0	238	6	ADA25195	Ada25195 Green flu
2	1272	99.7	238	6	ADA25221	Ada25221 Green flu
3	1270	99.5	238	6	ADA25216	Ada25216 Green flu
4	1268	99.4	238	6	ADA25218	Ada25218 Green flu
5	1268	99.4	238	6	ADA25222	Ada25222 Green flu
6	1266	99.2	238	5	AAB16087	Aae16087 Aequorea
7	1266	99.2	238	6	ADA25215	Ada25215 Green flu
8	1266	99.2	238	6	ADA25223	Ada25223 Green flu
9	1265	99.1	238	6	ADA25196	Ada25196 Green flu
10	1264	99.1	238	6	ADA25214	Ada25214 Green flu
11	1263	99.0	238	6	ABR44424	Abr44424 F64L-Y66H
12	1262	98.9	238	6	ADA25212	Ada25212 Green flu
13	1262	98.9	238	5	ADA25220	Ada25220 Green flu
14	1260	98.7	238	6	ABG32368	Abg32368 Aequorea
15	1260	98.7	238	6	AAE34996	Aae34996 Aequorea
16	1260	98.7	238	6	ADA25219	Ada25219 Green flu
17	1259	98.7	238	2	AAW76106	Aaw76106 A. victor
18	1259	98.7	238	5	AAE16073	Aae16073 Aequorea
19	1258	98.6	238	6	AAW05304	Aaw05304 Green flu
20	1258	98.6	238	2	AAW24232	Aaw24232 Aequorea
21	1258	98.6	238	2	AAW76105	Aaw76105 A. victor
22	1258	98.6	238	2	AAW40479	Aaw40479 A. victor
23	1258	98.6	238	2	AAW76371	Aaw76371 A. victor
24	1258	98.6	238	4	AAB73552	Aab73552 Wild-type
25	1258	98.6	238	5	AAE16038	Aae16038 Aequorea
26	1258	98.6	238	5	ABG32365	Abg32365 Aequorea
27	1258	98.6	238	6	ABG75980	Abg75980 Jellyfish
28	1258	98.6	238	6	AAB34999	Aae34999 Aequorea
29	1258	98.6	238	6	AAB34985	Aae34985 Wild-type
30	1258	98.6	238	6	ABR44423	Abr44423 Green flu
31	1258	98.6	238	6	ADA25217	Ada25217 Green flu
32	1258	98.6	238	6	ADA25194	Ada25194 Aequorea
33	1258	98.6	238	7	ADF70380	Adf70380 Aequorea
34	1258	98.6	238	7	ADM78505	Adm78505 Wild-type
35	1258	98.6	238	7	ADM78577	Adm78577 Mutant Ae
36	1258	98.6	238	8	ADQ59552	Adq59552 Aequorea
37	1257	98.5	238	8	AAW65084	Aaw65084 A. victor
38	1257	98.5	238	5	ABG32367	Abg32367 Aequorea
39	1257	98.5	238	6	ABR44425	Abr44425 F64L-Y66H
40	1255	98.4	238	6	AAW22101	Aaw22101 Aequorea
41	1255	98.4	238	2	AAW76111	Aaw76111 A. victor
42	1255	98.4	238	2	AAW76110	Aaw76110 A. victor
43	1255	98.4	238	2	AAW76109	Aaw76109 A. victor
44	1255	98.4	238	2	AAW65079	Aaw65079 A. victor
45	1255	98.4	238	2	AAW65078	Aaw65078 A. victor

ALIGNMENTS

```
RESULT 1
ADA25195
ID   ADA25195 standard; protein; 238 AA.
XX
AC   ADA25195;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Green fluorescent protein mutant, F64L-S175G-E222G-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key            Location/Qualifiers
FT   Misc-difference 64
FT                  /note= "Leu replaces wild-type Phe"
FT   Misc-difference 175
FT                  /note= "Gly replaces wild-type Ser"
FT   Misc-difference 222
FT                  /note= "Gly replaces wild-type Glu"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-00023288.
XX
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AE, Michael NP, Thomas N;
XX
DR   WPI; 2003-095652/09.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Claim 8; Fig 3; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid subsitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proistion 65
```

CC or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC invention exhibit enhanced fluorescence relative to wild type GFP when
CC expressed in non-homologous cells at temperatures above 30 degrees
CC Celsius, and excited at 490 nm. The mutants can also be detected in
CC mammalian cells at lower levels of expression and with increased
CC sensitivity relative to wild type GFP. The GFP mutants of the invention
CC are useful as non-toxic markers for selection of transfected cells, as
CC protein labels in living and fixed cells, as markers in cell or organelle
CC fusion, for visualising translocation of intracellular proteins to a
CC specific organelle, as secretion markers, as genetic reporters or protein
CC tags for protein and gene expression in transgenic animals, as cell or
CC organelle integrity markers, as transfection markers, as markers to be
CC used in combination with fluorescent activated cell sorting (FACS), as
CC real-time probes working at near physiological concentrations, for
CC performing transposon vector mutagenesis, and as reporters for bacterial
CC detection. The present sequence represents the specifically claimed
CC Aequorea victoria GFP mutant F64L-S175G-E222G-GFP.
XX
SQ Sequence 238 AA;

Query Match   100.0%;   Score 1276; DB 6; Length 238;
Best Local Similarity 100.0%; Pred. No. 4.9e-125;
Matches 238; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
```

RESULT 2
ADA25221
ID ADA25221 standard; protein; 238 AA.
XX
AC ADA25221;
XX
DT 20-NOV-2003 (first entry)
XX
DE Green fluorescent protein mutant, S175G-E222G-GFP.
XX
KW Green fluorescent protein; GFP; jellyfish; marker protein;
KW reporter protein; mutant; mutein.
XX
OS Aequorea victoria.
XX
FH Key   Location/Qualifiers
FT Misc-difference 175
FT   /note= "Gly replaces wild-type Ser"
FT Misc-difference 222
FT   /note= "Gly replaces wild-type Glu"
XX
PN GB2374868-A.
XX
PD 30-OCT-2002.
XX
PF 28-SEP-2001; 2001GB-00023288.
XX
PR 23-APR-2001; 2001GB-00009858.
XX
PA (AMSH) AMERSHAM PHARMACIA BIOTECH UK LTD.

PA (AMSH) AMERSHAM BIOSCIENCES UK LTD.
XX
PI Stubbs SLJ, Jones AE, Michael NP, Thomas N;
XX
DR WPI; 2003-095652/09.
DR N-PSDB; ADA25193.
XX
PT Novel fluorescent protein derived from green fluorescent protein useful
PT as a transfection marker, has different excitation spectrum and/or
PT emission spectrum compared with wild-type green fluorescent protein.
XX
PS Example 2; Page; 52pp; English.
XX
CC The invention relates to Aequorea victoria green fluorescent protein
CC (GFP) mutants containing an amino acid substitution at positions 64 and
CC 175, and additionally an amino acid substitution at either proistion 65
CC or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC invention exhibit enhanced fluorescence relative to wild type GFP when
CC expressed in non-homologous cells at temperatures above 30 degrees
CC Celsius, and excited at 490 nm. The mutants can also be detected in
CC mammalian cells at lower levels of expression and with increased
CC sensitivity relative to wild type GFP. The GFP mutants of the invention
CC are useful as non-toxic markers for selection of transfected cells, as
CC protein labels in living and fixed cells, as markers in cell or organelle
CC fusion, for visualising translocation of intracellular proteins to a
CC specific organelle, as secretion markers, as genetic reporters or protein
CC tags for protein and gene expression in transgenic animals, as cell or
CC organelle integrity markers, as transfection markers, as markers to be
CC used in combination with fluorescent activated cell sorting (FACS), as
CC real-time probes working at near physiological concentrations, for
CC performing transposon vector mutagenesis, and as reporters for bacterial
CC detection. The present sequence represents an Aequorea victoria GFP
CC mutant used in an example of the invention. Note: The present sequence is
CC not shown in the specification, but is derived from the wild-type GFP
CC sequence shown in Fig 2 and the information given on page 24.
XX
SQ Sequence 238 AA;

Query Match   99.7%;   Score 1272; DB 6; Length 238;
Best Local Similarity 99.6%; Pred. No. 1.3e-124;
Matches 237; Conservative   0; Mismatches   1; Indels   1; Gaps   0;

```
Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTFSYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
```

RESULT 3
ADA25216
ID ADA25216 standard; protein; 238 AA.
XX
AC ADA25216;
XX
DT 20-NOV-2003 (first entry)
XX
DE Green fluorescent protein mutant, F64L-E222G-GFP.
XX
KW Green fluorescent protein; GFP; jellyfish; marker protein;
KW reporter protein; mutant; mutein.
XX

```
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 64
FT                   /note= "Leu replaces wild-type Phe"
FT   Misc-difference 222
FT                   /note= "Gly replaces wild-type Glu"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-00023288.
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AE, Michael NP, Thomas N;
XX
DR   WPI; 2003-095652/09.
DR   N-PSDB; ADA25193.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Example 2; Page; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid substitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proistion 65
CC   or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), as
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial
CC   detection. The present sequence represents an Aequorea victoria GFP
CC   mutant used in an example of the invention. Note: The present sequence is
CC   not shown in the specification, but is derived from the wild-type GFP
CC   sequence shown in Fig 2 and the information given on page 24.
XX
SQ   Sequence 238 AA;

Query Match             99.5%;   Score 1270;  DB 6;  Length 238;
Best Local Similarity   99.6%;   Pred No. 2.1e-124;
Matches  237;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKPICTTGKLPVPWPTL  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKPICTTGKLPVPWPTL  60

Qy      61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy     121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
```

```
Qy     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238

RESULT 4
ADA25218
ID   ADA25218 standard; protein; 238 AA.
XX
AC   ADA25218;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Green fluorescent protein mutant, F64L-S175G-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 64
FT                   /note= "Leu replaces wild-type Phe"
FT   Misc-difference 175
FT                   /note= "Gly replaces wild-type Ser"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-00023288.
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AE, Michael NP, Thomas N;
XX
DR   WPI; 2003-095652/09.
DR   N-PSDB; ADA25193.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Example 2; Page; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid substitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proistion 65
CC   or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), as
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial
CC   detection. The present sequence represents an Aequorea victoria GFP
CC   mutant used in an example of the invention. Note: The present sequence is
CC   not shown in the specification, but is derived from the wild-type GFP
CC   sequence shown in Fig 2 and the information given on page 24.
XX
```

```
'SQ   Sequence 238 AA;

Query Match          99.4%;     Score 1268;   DB 6;   Length 238;
Best Local Similarity 99.6%;    Pred. No. 3.4e-124;
Matches 237;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy    61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
           |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 5
ADA25222
ID   ADA25222 standard; protein; 238 AA.
XX
AC   ADA25222;
XX
DT   20-NOV-2003   (first entry)
XX
DE   Green fluorescent protein mutant, V163A-S175G-E222G-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 163
FT                   /note= "Ala replaces wild-type Val"
FT   Misc-difference 175
FT                   /note= "Gly replaces wild-type Ser"
FT   Misc-difference 222
FT                   /note= "Gly replaces wild-type Glu"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-00023288.
XX
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AE, Michael NP, Thomas N;
XX
DR   WPI; 2003-095652/09.
DR   N-PSDB; ADA25193.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Example 2; Page; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid substitution at positions 64 and
CC   175, and additionally an amino acid substitution at either position 65
CC   or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), as
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial
CC   detection. The present sequence represents an Aequorea victoria GFP
CC   mutant used in an example of the invention. Note: The present sequence is
CC   not shown in the specification, but is derived from the wild-type GFP
CC   sequence shown in Fig 2 and the information given on page 24.
XX
SQ   Sequence 238 AA;

Query Match          99.4%;     Score 1268;   DB 6;   Length 238;
Best Local Similarity 99.2%;    Pred. No. 3.4e-124;
Matches 236;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy    61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKIRHNIEDGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238


RESULT 6
AAE16087
ID   AAE16087 standard; protein; 238 AA.
XX
AC   AAE16087;
XX
DT   26-MAR-2002   (first entry)
XX
DE   Aequorea victoria GFP mutant protein (E222G).
XX
KW   Green fluorescent protein; cell lineage tracer; protein localisation;
KW   GFP; fusion tag; gene expression marker; fluorescent energy acceptor;
KW   immunoassay; hybridisation assay; fluorescent energy donor; biosensor;
KW   FRET; fluorescence resonance energy transfer; mutant; mutein.
XX
OS   Aequorea victoria.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 222
FT                   /note= "Wild type Glu substituted with Gly"
XX
PN   WO200190147-A2.
XX
PD   29-NOV-2001.
XX
PF   17-MAY-2001; 2001WO-US016149.
XX
PR   19-MAY-2000; 2000US-00575847.
XX
```

```
PA   (UYOR-) UNIV OREGON STATE.
XX
PI   Wachter R,  Remington SJ;
XX
DR   WPI; 2002-083084/11.
XX
PT   New long wavelength engineered fluorescent proteins, useful as markers
PT   for gene expression, tracers of cell lineage or as fusion tags to monitor
PT   protein localization, or in detection assays, e.g. immunoassays or
PT   hybridization assays.
XX
PS   Example; Page; 181pp; English.
XX
CC   The patent discloses functional engineered fluorescent protein and its
CC   corresponding polynucleotide. The amino acid sequence of the engineered
CC   protein is identical to Aequorea green fluorescent protein (GFP). The
CC   engineered fluorescent proteins of the invention have varied fluorescent
CC   properties and have the ability to respond to ion concentrations via a
CC   change in fluorescent characteristics. They are useful as markers for
CC   gene expression, tracers of cell lineage or as fusion tags to monitor
CC   protein localisation within living cells. The engineered fluorescent
CC   proteins are particularly useful for coupling engineered fluorescent
CC   proteins to antibodies, nucleic acids or other receptors for use in
CC   detection assays, e.g. immunoassays or hybridisation assays. They are
CC   useful for tracking the movement of proteins in cells or in systems for
CC   detecting induction of transcription and for the simultaneous measurement
CC   of two or more processes within cells. Proteins of the invention are also
CC   useful as fluorescent energy donors or acceptors as well as biosensors
CC   for detecting anions. They are also useful in fluorescence resonance
CC   energy transfer (FRET). The crystal structure of the GFP is useful for
CC   designing mutants having altered fluorescent characteristics which are
CC   particularly useful to identify amino acids whose substitution alters the
CC   fluorescent properties of the protein. The crystal structure of the GFP
CC   is also useful for designing mutants having altered anion binding
CC   characteristics which are particularly useful for identifying amino acids
CC   whose substitution alters the specificity and affinity of the binding
CC   site to various anions, and for monitoring anion binding and therefore
CC   the concentration of the anion. The present sequence is Aequorea victoria
CC   GFP mutant protein (E222G). Note: This sequence is not shown in the
CC   specification but is derived from Aequorea victoria wild type green
CC   fluorescent protein shown in figure 3 of the specification (AAB16038)
XX
SQ   Sequence 238 AA;

Query Match         99.2%;  Score 1266;  DB 5;  Length 238;
Best Local Similarity 99.2%;  Pred. No. 5.5e-124;
Matches 236;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIBLKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGPVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGPVTAAGITHGMDELYK 238

RESULT 7
ADA25215
ID   ADA25215 standard; protein; 238 AA.
XX
AC   ADA25215;
XX
DT   20-NOV-2003 (first entry)
XX
```

```
DB   Green fluorescent protein mutant, E222G-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 222
FT                   /note= "Gly replaces wild-type Glu"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-00023288.
XX
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ,  Jones AE,  Michael NP,  Thomas N;
XX
DR   WPI; 2003-095652/09.
DR   N-PSDB; ADA25193.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Example 2; Page; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid substitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proistion 65
CC   or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), in
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial
CC   detection. The present sequence represents an Aequorea victoria GFP
CC   mutant used in an example of the invention. Note: The present sequence is
CC   not shown in the specification, but is derived from the wild-type GFP
CC   sequence shown in Fig 2 and the information given on page 24.
XX
SQ   Sequence 238 AA;

Query Match         99.2%;  Score 1266;  DB 6;  Length 238;
Best Local Similarity 99.2%;  Pred. No. 5.5e-124;
Matches 236;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIBLKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
```

```
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
             |||||||||||||||||||||||||||||||||||||||||||| ||||||
Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238


RESULT 8
ADA25223
ID   ADA25223 standard; protein; 238 AA.
XX
AC   ADA25223;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Green fluorescent protein mutant, F64L-V163A-E222G-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 64
FT                   /note= "Leu replaces wild-type Phe"
FT   Misc-difference 163
FT                   /note= "Ala replaces wild-type Val"
FT   Misc-difference 222
FT                   /note= "Gly replaces wild-type Glu"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-00023288.
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AE, Michael NP, Thomas N;
XX
DR   WPI; 2003-095652/09.
DR   N-PSDB; ADA25193.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Example 2; Page 2; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid subsitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proistion 65
CC   or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), as
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial
CC   detection. The present sequence represents an Aequorea victoria GFP
CC   mutant used in an example of the invention. Note: The present sequence is
CC   not shown in the specification, but is derived from the wild-type GFP
XX   sequence shown in Fig 2 and the information given on page 24.
SQ   Sequence 238 AA;

Query Match             99.2%;  Score 1266;  DB 6;  Length 238;
Best Local Similarity   99.2%;  Pred. No. 5.5e-124;
Matches  236;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy    61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKIRHNIEDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238


RESULT 9
ADA25196
ID   ADA25196 standard; protein; 238 AA.
XX
AC   ADA25196;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Green fluorescent protein mutant, F64L-S65T-S175G-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 64
FT                   /note= "Leu replaces wild-type Phe"
FT   Active-site     65
FT                   /note= "Thr replaces wild-type Ser"
FT   Misc-difference 175
FT                   /note= "Gly replaces wild-type Ser"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-00023288.
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AE, Michael NP, Thomas N;
XX
DR   WPI; 2003-095652/09.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Claim 9; Fig 4; 52pp; English.
XX
```

CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid subsitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proistion 65
CC   or 222. The mutants of the invention are particularly F64L-S175G-
CC   GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), as
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial
CC   detection. The present sequence represents the specifically claimed
CC   Aequorea victoria GFP mutant F64L-S65T-S175G-GFP.
XX
SQ   Sequence 238 AA;

Query Match   99.1%; Score 1265; DB 6; Length 238;
Best Local Similarity   99.2%; Pred. No. 7e-124;
Matches   236; Conservative   1; Mismatches   1; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238

RESULT 10
ADA25214
ID   ADA25214 standard; protein; 238 AA.
XX
AC   ADA25214;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Green fluorescent protein mutant, S175G-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 175
FT   /note= "Gly replaces wild-type Ser"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-0023288.
XX
PR   23-APR-2001; 2001GB-00009858.
XX

PA   (AMSH) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AE, Michael NP, Thomas N;
XX
DR   WPI; 2003-095652/09.
DR   N-PSDB; ADA25193.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Example 2; Page; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid subsitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proistion 65
CC   or 222. The mutants of the invention are particularly F64L-S175G-
CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), as
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial
CC   detection. The present sequence represents an Aequorea victoria GFP
CC   mutant used in an example of the invention. Note: The present sequence is
CC   not shown in the specification, but is derived from the wild-type GFP
CC   sequence shown in Fig 2 and the information given on page 24.
XX
SQ   Sequence 238 AA;

Query Match   99.1%; Score 1264; DB 6; Length 238;
Best Local Similarity   99.2%; Pred. No. 8.9e-124;
Matches   236; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              |||·||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              |||·||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NRIBLKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||||||||||||||||||·|||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 11
ABR44424
ID   ABR44424 standard; protein; 238 AA.
XX
AC   ABR44424;
XX
DT   11-JUL-2003 (first entry)
XX
DE   F64L-Y66H-S175G-GFP #SEQ ID 3.
XX
KW   Green fluorescent protein; GFP; marker; label; reporter;
KW   bacterial detection; mutein.
XX

```
XX                                                    AC   ADA25212;
OS   Aequorea victoria.                               XX
XX                                                    DT   20-NOV-2003  (first entry)
PN   WO200302929286-A1.                               XX
XX                                                    DE   Green fluorescent protein mutant, F64L-GFP.
PD   10-APR-2003.                                     XX
XX                                                    KW   Green fluorescent protein; GFP; jellyfish; marker protein;
PF   27-SEP-2002; 2002WO-GB004354.                    KW   reporter protein; mutant; mutein.
XX                                                    XX
PR   28-SEP-2001; 2001GB-00023314.                    OS   Synthetic.
XX                                                    OS   Aequorea victoria.
PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.             XX
PA   (STUB/) STUBBS S.                                FH   Key             Location/Qualifiers
XX                                                    FT   Misc-difference 64
PI   Jones AE, Davies J, Ruehlmann D, Michael NP;     FT                   /note= "Leu replaces wild-type Phe"
XX                                                    XX
DR   WPI; 2003-421211/39.                             PN   GB2374868-A.
XX                                                    XX
PT   Novel blue-shifted variants of green             PD   30-OCT-2002.
PT   fluorescent protein having improved              XX
PT   fluorescence properties, useful for measuring    PF   28-SEP-2001; 2001GB-00023288.
PT   expression of a protein of interest in a cell,   XX
PT   as a secretion, cell or organelle marker.        PR   23-APR-2001; 2001GB-00009858.
XX                                                    XX
PS   Claim 7; Fig 3; 45pp; English.                   PA   (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
XX                                                    PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
CC   The invention relates to a fluorescent protein   XX
CC   derived from Green fluorescent Protein (GFP),    PI   Stubbs SLJ, Jones AE, Michael NP, Thomas N;
CC   or any functional GFP analog having an amino     XX
CC   acid sequence which is modified by amino acid    DR   WPI; 2003-095652/09.
CC   substitution compared with the sequence of       DR   N-PSDB; ADA25193.
CC   wild type (wt) GFP. The polynucleotide encoding  XX
CC   a GFP of the invention is useful for measuring   PT   Novel fluorescent protein derived from green
CC   the expression of a protein of interest in a     PT   fluorescent protein useful as a transfection
CC   cell, and is also useful for determining the     PT   marker, has different excitation spectrum and/or
CC   cellular and/or extracellular localisation of a  PT   emission spectrum compared with wild-type green
CC   protein of interest. GFP's of the invention      PT   fluorescent protein.
CC   are useful as non-toxic markers for selection    XX
CC   of transfected cells containing an expression    PS   Example 2; Page; 52pp; English.
CC   vector encoding at least the fluorescent         XX
CC   protein. GFP's may also be used as protein       CC   The invention relates to Aequorea victoria green
CC   labels in living and fixed cells, as markers     CC   fluorescent protein (GFP) mutants containing an
CC   in cell or organelle fusion, for visualising     CC   amino acid substitution at positions 64 and
CC   translocation of proteins fused to them, as      CC   175, and additionally an amino acid substitution
CC   secretion markers, as genetic reporters or       CC   at either proistion 65 or 222. The mutants of
CC   protein tags in transgenic animals, as cell or   CC   the invention are particularly F64L-S175G-E222G-
CC   organelle integrity markers, as markers to be    CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196).
CC   used in combination with fluorescent             CC   GFP mutants of the invention exhibit enhanced
CC   activated cell sorting (FACS) and as reporters   CC   fluorescence relative to wild type GFP when
CC   for bacterial detection. GFP's are also useful   CC   expressed in non-homologous cells at temperatures
CC   for performing transposon mutagenesis, where     CC   above 30 degrees Celsius, and excited at 490 nm.
CC   the GFP is used as a marker in transcriptional   CC   The mutants can also be detected in mammalian
CC   and translational fusions. GFP's of the          CC   cells at lower levels of expression and with
CC   invention have improved fluorescence properties  CC   increased sensitivity relative to wild type GFP.
CC   and are suitable labels for proteins present at  CC   The GFP mutants of the invention are useful as
CC   low concentrations. The current sequence         CC   non-toxic markers for selection of transfected
CC   represents the green fluorescent protein (GFP)   CC   cells, as protein labels in living and fixed
CC   mutant F64L-Y66H-S175G-GFP                       CC   cells, as markers in cell or organelle fusion,
XX                                                    CC   for visualising translocation of intracellular
SQ   Sequence 238 AA;                                 CC   proteins to a specific organelle, as secretion
                                                      CC   markers, as genetic reporters or protein tags
                                                      CC   for protein and gene expression in transgenic
                                                      CC   animals, as cell or organelle integrity markers,
                                                      CC   as transfection markers, as markers to be used
                                                      CC   in combination with fluorescent activated cell
                                                      CC   sorting (FACS), as real-time probes working at
                                                      CC   near physiological concentrations, for
                                                      CC   performing transposon vector mutagenesis, and
                                                      CC   as reporters for bacterial detection. The present
                                                      CC   sequence represents an Aequorea victoria GFP
                                                      CC   mutant used in an example of the invention. Note:
                                                      CC   The present sequence is not shown in the
                                                      CC   specification, but is derived from the wild-type GFP
                                                      CC   sequence shown in Fig 2 and the information given on page 24.
                                                      XX
                                                      SQ   Sequence 238 AA;

Query Match             99.0%;  Score 1263;  DB 6;    Query Match             98.9%;  Score 1262;  DB 6;  Length 238;
Length 238;                                           Best Local Similarity   99.2%;  Pred. No. 1.4e-123;
Best Local Similarity   99.2%;  Pred. No. 1.1e-123;     Matches  236;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;
  Matches  236;  Conservative   1;  Mismatches   1;
  Indels   0;  Gaps   0;                              Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
                                                                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATY  Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              GKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||||||||||||||||||          Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATY
              GKLTLKFICTTGKLPVPWPTL 60

Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIF
              FKDDGNYKTRAEVKFEGDTLV 120
              |||||:|||||||||||||||||||||||||||||||||
Db         61 VTTLSHGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIF
              FKDDGNYKTRAEVKFEGDTLV 120

Qy        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKN
              GIKVNPKIRHNIEDGGVQLAD 180
              ||||||||||||||||||||||||||||||||||||||||
Db        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKN
              GIKVNFKIRHNIEDGGVQLAD 180

Qy        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMV
              LLGFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||||||||||||||||
Db        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMV
              LLEFVTAAGITHGMDELYK 238

RESULT 12
ADA25212
ID   ADA25212 standard; protein; 238 AA.
XX
```

```
Db    61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKIRHNIEDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
         ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLBFVTAAGITHGMDELYK 238

RESULT 13
ADA25220
ID    ADA25220 standard; protein; 238 AA.
XX
AC    ADA25220;
XX
DT    20-NOV-2003  (first entry)
XX
DE    Green fluorescent protein mutant, V163A-E222G-GFP.
XX
KW    Green fluorescent protein; GFP; jellyfish; marker protein;
KW    reporter protein; mutant; mutein.
XX
OS    Synthetic.
OS    Aequorea victoria.
XX
FH    Key              Location/Qualifiers
FT    Misc-difference  163
FT                     /note= "Ala replaces wild-type Val"
FT    Misc-difference  222
FT                     /note= "Gly replaces wild-type Glu"
XX
PN    GB2374868-A.
XX
PD    30-OCT-2002.
XX
PP    28-SEP-2001;  2001GB-00023288.
XX
PR    23-APR-2001;  2001GB-00009858.
XX
PA    (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA    (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
XX
PI    Stubbs SLJ,  Jones AE,  Michael NP,  Thomas N;
XX
DR    WPI; 2003-095652/09.
DR    N-PSDB; ADA25193.
XX
PT    Novel fluorescent protein derived from green fluorescent protein useful
PT    as a transfection marker, has different excitation spectrum and/or
PT    emission spectrum compared with wild-type green fluorescent protein.
XX
PS    Example 2; Page; 52pp; English.
XX
CC    The invention relates to Aequorea victoria green fluorescent protein
CC    (GFP) mutants containing an amino acid substitution at positions 64 and
CC    175, and additionally an amino acid substitution at either proistion 65
CC    or 222. The mutants of the invention are particularly F64L-S175G-GFP
CC    GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC    invention exhibit enhanced fluorescence relative to wild type GFP when
CC    expressed in non-homologous cells at temperatures above 30 degrees
CC    Celsius, and excited at 490 nm. The mutants can also be detected in
CC    mammalian cells at lower levels of expression and with increased
CC    sensitivity relative to wild type GFP. The GFP mutants of the invention
CC    are useful as non-toxic markers for selection of transfected cells, as
CC    protein labels in living and fixed cells, as markers in cell or organelle
CC    fusion, for visualising translocation of intracellular proteins to a
CC    specific organelle, as secretion markers, as genetic reporters or protein
CC    tags for gene expression in transgenic animals, as cell or
CC    organelle integrity markers, as transfection markers, as markers to be
CC    used in combination with fluorescent activated cell sorting (FACS), as
```

```
CC    real-time probes working at near physiological concentrations, for
CC    performing transposon vector mutagenesis, and as reporters for bacterial
CC    detection. The present sequence represents an Aequorea victoria GFP
CC    mutant used in an example of the invention. Note: The present sequence is
CC    not shown in the specification, but is derived from the wild-type GFP
CC    sequence shown in Fig 2 and the information given on page 24.
XX
SQ    Sequence 238 AA;

Query Match             98.9%;  Score 1262;  DB 6;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 1.4e-123;
Matches 235;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy    61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKIRHNIEDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238

RESULT 14
ABG32368
ID    ABG32368 standard; protein; 238 AA.
XX
AC    ABG32368;
XX
DT    15-NOV-2002  (first entry)
XX
DE    Aequorea victoria green fluorescent protein (GFP), mutant F64M.
XX
KW    Green fluorescent protein; GFP; jellyfish; chromophore development;
KW    bioluminescent labeling; agriculture; nitrogen fixation; cheese;
KW    food production; anaerobic fermentation; alcoholic fermentation; yogurt;
KW    pickle; beer; methanol; fuel production; glycerol; sewage treatment;
KW    soil; detoxifying microbial agent; mutant; mutein.
XX
OS    Aequorea victoria.
OS    Synthetic.
XX
FH    Key              Location/Qualifiers
FT    Misc-difference  64
FT                     /note= "wild type Phe substituted by Met"
XX
PN    US641411-B1.
XX
PD    02-JUL-2002.
XX
PP    15-OCT-1999;  99US-00418785.
XX
PR    16-OCT-1998;  98US-0104563P.
XX
PA    (RUTF ) UNIV RUTGERS STATE NEW JERSEY.
XX
PI    Fisher H;
XX
DR    WPI; 2002-641578/69.
XX
PT    Novel green fluorescent protein variant useful in bioluminescent
PT    labeling, e.g. as in situ marker of low oxygen status in cells and
PT    tissues, exhibits rapid chromophore development under low oxygen
PT    concentration.
XX
PS    Claim 1; Page; 18pp; English.
```

XX
PI Tsien RY, Ting AY, Zhang J;
DR WPI; 2003-148474/14.
XX
PT Novel chimeric phosphorylation indicators, useful for detecting
PT kinase/phosphatase in samples, has donor molecule, phosphorylatable
PT domain, phosphoaminoacid binding domain, and acceptor molecule, in
PT operative linkage.
XX
PS Disclosure; Col; 38pp; English.
XX
CC The present invention relates to chimeric phosphorylation indicators
CC comprising a phosphorylation polypeptide and a fluorescent protein or in
CC operative linkage, a donor molecule, a phosphorylatable domain, a
CC phosphoaminoacid binding domain (PAABD) and an acceptor molecule. The
CC phosphorylation indicators of the invention are useful for detecting
CC kinases or phosphatases in a biological sample. They are also useful in
CC high throughput analysis e.g. for detecting a kinase inhibitor or
CC phosphatase inhibitor. The present sequence is Aequorea victoria mutant
CC green fluorescent protein (GFP; S175G) used in the invention. Note: This
CC sequence is not shown in the specification but is derived from Aequorea
CC victoria wild-type GFP shown as SEQ ID NO: 2 in column 53-54 of the
CC specification (AAB34957)
XX
SQ Sequence 238 AA;

Query Match   98.7%; Score 1260; DB 6; Length 238;
Best Local Similarity 98.7%; Pred. No. 2.3e-123;
Matches 235; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

Search completed: September 16, 2005, 17:16:10
Job time : 119 secs

XX
CC The invention relates to a green fluorescent protein (GFP) variant (I)
CC exhibiting, under oxygen concentration that is equal or lower than
CC atmospheric, rapid chromophore development as compared with a wild-type
CC GFP. (I) is useful in in vitro, in situ or in vivo bioluminescent
CC labeling, in any labeling applications in which GFP is currently used, in
CC situations where wild-type GFP is not stable i.e., high ionic strength
CC buffers and/or high temperature, in haplophytes, thermophytes and in
CC organelles with varying salt and membrane compositions, and in
CC agriculture e.g. in nitrogen fixation. (I) is useful as an in situ marker
CC of low oxygen status in cells and tissues, and as detectable substance in
CC any application in food production where anaerobic or alcoholic
CC fermentation is utilised, including manufacturing of cheese, yogurt,
CC pickles, beer and alcohol fermentations other than ethanol such as
CC methanol for fuel production or glycerol. (I) is useful for monitoring
CC oxygen level in microbial degradation of municipal sewage treatment, for
CC testing an enzyme having a degradative capacity under low oxygen
CC condition, for monitoring microbial processes in soil, and for monitoring
CC activity of detoxifying microbial agents. The present sequence represents
CC the amino acid sequence of an Aequorea victoria GFP mutant protein
XX
SQ Sequence 238 AA;

Query Match   98.7%; Score 1260; DB 5; Length 238;
Best Local Similarity 98.7%; Pred. No. 2.3e-123;
Matches 235; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTMSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 15
AAB34996
ID AAB34996 standard; protein; 238 AA.
XX
AC AAB34996;
XX
DT 28-MAY-2003 (first entry)
XX
DE Aequorea victoria mutant green fluorescent protein (S175G).
XX
KW Phosphorylation indicator; fluorescent protein; detection; phosphatase;
KW kinase; green fluorescent protein; GFP; mutant; mutein.
XX
OS Aequorea victoria.
OS Synthetic.
XX
FH Key   Location/Qualifiers
FT Misc-difference 175
FT   /note= "Wild-type Ser substituted with Gly"
XX
PN WO200295058-A2.
XX
PD 28-NOV-2002.
XX
PF 24-MAY-2002; 2002WO-US016955.
XX
PR 24-MAY-2001; 2001US-00865291.
XX
PA (REGC) UNIV CALIFORNIA.

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      September 16, 2005, 17:08:27 ; Search time 31 Seconds
                              (without alignments)
                           573.112 Million cell updates/sec

Title:       US-10-757-624-3
Perfect score: 1276
Sequence:    1 MSKGEELFTGVVPILVELDG.........VLLGFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62
             Gapop 10.0 , Gapext 0.5

Searched:    513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:        513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued Patents AA:*
             1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
             2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
             3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
             4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
             5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep:*
             6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Match Length	DB	ID	Description	
1	1258	98.6	238	1	US-08-753-143-2	Sequence 2, Appli
2	1258	98.6	238	2	US-08-679-865-2	Sequence 2, Appli
3	1258	98.6	238	2	US-08-680-876-2	Sequence 2, Appli
4	1258	98.6	238	2	US-08-792-553-2	Sequence 2, Appli
5	1258	98.6	238	3	US-08-753-144-2	Sequence 2, Appli
6	1258	98.6	238	3	US-09-094-359-2	Sequence 2, Appli
7	1258	98.6	238	3	US-09-172-063-2	Sequence 2, Appli
8	1258	98.6	238	3	US-09-263-975-2	Sequence 2, Appli
9	1258	98.6	238	3	US-08-727-452-2	Sequence 2, Appli
10	1258	98.6	238	4	US-09-418-785-1	Sequence 1, Appli
11	1258	98.6	238	4	US-09-129-192C-2	Sequence 74, Appl
12	1258	98.6	238	4	US-09-129-192C-74	Sequence 2, Appli
13	1258	98.6	238	4	US-09-602-641-2	Sequence 2, Appli
14	1258	98.6	238	4	US-09-704-463-2	Sequence 2, Appli
15	1258	98.6	238	4	US-10-024-686A-2	Sequence 2, Appli
16	1258	98.6	238	4	US-08-594-575C-2	Sequence 16, Appl
17	1255	98.4	238	3	US-08-893-327-16	Sequence 18, Appl
18	1255	98.4	247	3	US-08-893-327-18	Sequence 2, Appli
19	1254	98.3	238	1	US-08-337-915A-2	Sequence 1, Appli
20	1254	98.3	238	3	US-09-121-539-1	Sequence 2, Appli
21	1254	98.3	238	4	US-09-214-909-2	Sequence 10, Appl
22	1254	98.3	238	4	US-09-479-645A-10	Sequence 159, App
23	1254	98.3	238	4	US-09-479-645A-159	Sequence 4, Appli
24	1254	98.3	238	4	US-09-472-065A-4	Sequence 4, Appli
25	1254	98.3	238	5	PCT-US95-14692-2	Sequence 2, Appli
26	1254	98.3	238	4	US-09-472-065A-6	Sequence 6, Appli
27	1253	98.2	238		US-09-023-946B-35	Sequence 35, Appl
28	1253	98.2	238	4	US-08-646-538-2	Sequence 2, Appli
29	1253	98.2	239	3	US-09-503-222-2	Sequence 2, Appli
30	1253	98.2	239	3	US-09-472-065A-2	Sequence 2, Appli
31	1251	98.0	238	4	US-09-479-645A-94	Sequence 94, Appl
32	1251	98.0	243	4	US-09-479-645A-96	Sequence 96, Appl
33	1251	98.0	243	4	US-09-479-645A-98	Sequence 98, Appl
34	1251	98.0	243	4	US-09-479-645A-100	Sequence 100, App
35	1251	98.0	243	4	US-09-479-645A-102	Sequence 102, App
36	1251	98.0	243	4	US-09-479-645A-104	Sequence 104, App
37	1251	98.0	243	4	US-09-479-645A-110	Sequence 110, App
38	1251	98.0	243	4	US-09-863-901-1	Sequence 1, Appli
39	1251	98.0	432	4	US-09-863-901-2	Sequence 2, Appli
40	1251	98.0	441	4	US-09-863-901-3	Sequence 3, Appli
41	1251	98.0	450	4	US-09-863-901-4	Sequence 4, Appli
42	1251	98.0	468	4	US-09-863-901-5	Sequence 5, Appli
43	1251	98.0	477	4	US-09-863-901-6	Sequence 6, Appli
44	1251	98.0	906	4	US-09-091-042A-2	Sequence 2, Appli
45	1251	98.0	1070	4		

ALIGNMENTS

```
RESULT 1
US-08-753-143-2
; Sequence 2, Application US/08753143A
; GENERAL INFORMATION:
;  APPLICANT: Tsien, Roger Y.
;  APPLICANT: Heim, Roger
;  TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
;  FILE REFERENCE: 07257/032003
;  CURRENT APPLICATION NUMBER: US/08/753,143A
;  CURRENT FILING DATE: 1996-11-20
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;  LENGTH: 238
;  TYPE: PRT
;  ORGANISM: Aequorea victoria
US-08-753-143-2
```

```
Query Match             98.6%;  Score 1258;  DB 1;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 8.4e-129;
Matches  235;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||||.|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||||||||||||||||||.||||||||||||||||
Db        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

```
RESULT 2
US-08-679-865-2
; Sequence 2, Application US/08679865
; Patent No. 5912137
; GENERAL INFORMATION:
;  APPLICANT: Tsien, Roger Y.
;  APPLICANT: Cubitt, Andrew B.
;  TITLE OF INVENTION: Assays for Protein Kinases Using
;  TITLE OF INVENTION: Fluorescent Protein Substrates
;  NUMBER OF SEQUENCES: 48
;  CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/679,865
FILING DATE: 16-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John S.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 02307Z-069000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-679-865-2

Query Match 98.6%; Score 1258; DB 2; Length 238;
Best Local Similarity 98.7%; Pred. No. 8.4e-129;
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
        ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 3
US-08-680-876-2
; Sequence 2, Application US/08680876
; Patent No. 5925558
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Cubitt, Andrew B.
TITLE OF INVENTION: Assays for Protein Kinases Using
TITLE OF INVENTION: Fluorescent Protein Substrates
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,876
FILING DATE: 16-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John S.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 02307Z-069200
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-680-876-2

Query Match 98.6%; Score 1258; DB 2; Length 238;
Best Local Similarity 98.7%; Pred. No. 8.4e-129;
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
        ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 4
US-08-792-553-2
; Sequence 2, Application US/08792553
; Patent No. 5981200
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Heim, Roger
TITLE OF INVENTION: Tandem Fluorescent Protein Constructs
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & RICHARDSON P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/792,553
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lisa A. Haile, Ph.D.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/041001/UC 96-160-2
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 2:

```
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 238 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;       MOLECULE TYPE: protein
US-08-792-553-2

Query Match             98.6%;   Score 1258; DB 2; Length 238;
Best Local Similarity   98.7%;   Pred. No. 8.4e-129;
Matches  235; Conservative    0; Mismatches   3; Indels   0; Gaps   0;

Qy         1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db         1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy        61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy       121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy       181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
             |||||||||||||||||||||||||||||||||||| ||||||||||||
Db       181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 5
US-08-753-144-2
;  Sequence 2, Application US/08753144
;  Patent No. 6066476
;  GENERAL INFORMATION:
;     APPLICANT: Tsien, Roger Y.
;     APPLICANT: Heim, Roger
;     TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
;     NUMBER OF SEQUENCES: 5
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: Fish & Richardson P.C.
;        STREET: 4225 Executive Square, Suite 1400
;        CITY: La Jolla
;        STATE: CA
;        COUNTRY: USA
;        ZIP: 92037
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Diskette
;        COMPUTER: IBM Compatible
;        OPERATING SYSTEM: Windiwe95
;        SOFTWARE: FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/753,144
;        FILING DATE: 20-NOV-1996
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: 08/727,452
;        FILING DATE: 10-OCT-1996
;        APPLICATION NUMBER: US95/14692
;        FILING DATE: 13-NOV-1995
;        APPLICATION NUMBER: 08/337,915
;        FILING DATE: 10-NOV-1994
;     ATTORNEY/AGENT INFORMATION:
;        NAME: Haile, Lisa A.
;        REGISTRATION NUMBER: 38,347
;        REFERENCE/DOCKET NUMBER: 07257/032002
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE: 619/678-5070
;        TELEFAX: 619/678-5099
;  INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 238 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       FRAGMENT TYPE: internal
```

```
US-08-753-144-2

Query Match             98.6%;   Score 1258; DB 3; Length 238;
Best Local Similarity   98.7%;   Pred. No. 8.4e-129;
Matches  235; Conservative    0; Mismatches   3; Indels   3; Gaps   0;

Qy         1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db         1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy        61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy       121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy       181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
             |||||||||||||||||||||||||||||||||||| ||||||||||||
Db       181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 6
US-09-094-359-2
;  Sequence 2, Application US/09094359
;  Patent No. 6140132
;  GENERAL INFORMATION:
;     APPLICANT: Tsien, Roger Y.
;     APPLICANT: Miyawaki, Atsushi
;     APPLICANT: Llopis, Juan
;     TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
;     TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
;     FILE REFERENCE: 07257/067001
;     CURRENT APPLICATION NUMBER: US/09/094,359
;     CURRENT FILING DATE: 1998-06-09
;     NUMBER OF SEQ ID NOS: 18
;     SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 2
;       LENGTH: 238
;       TYPE: PRT
;       ORGANISM: Aequorea victoria
US-09-094-359-2

Query Match             98.6%;   Score 1258; DB 3; Length 238;
Best Local Similarity   98.7%;   Pred. No. 8.4e-129;
Matches  235; Conservative    0; Mismatches   3; Indels   3; Gaps   0;

Qy         1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db         1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy        61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy       121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy       181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
             |||||||||||||||||||||||||||||||||||| ||||||||||||
Db       181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 7
US-09-172-063-2
;  Sequence 2, Application US/09172063
;  Patent No. 6150176
;  GENERAL INFORMATION:
;     APPLICANT: Tsien, Roger Y.
;     APPLICANT: Miyawaki, Atsushi
;     APPLICANT: Llopis, Juan
```

```
;   APPLICANT: Wachter, Rebekka M.
;   APPLICANT: Remington, S. James
;   TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
;   TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
;   FILE REFERENCE: 07257/071001
;   CURRENT APPLICATION NUMBER: US/09/172,063
;   CURRENT FILING DATE: 1998-10-13
;   EARLIER APPLICATION NUMBER: 09/094,359
;   EARLIER FILING DATE: 1998-06-09
;   NUMBER OF SEQ ID NOS: 38
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 2
;     LENGTH: 238
;     TYPE: PRT
;     ORGANISM: Aequorea victoria
US-09-172-063-2

Query Match             98.6%;  Score 1258;  DB 3;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 8.4e-129;
Matches  235;           Conservative    0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 8
US-09-263-975-2
;   Sequence 2, Application US/09263975
;   Patent No. 6248550
;   GENERAL INFORMATION:
;   APPLICANT: Tsien, Roger Y.
;   APPLICANT: Cubitt, Andrew B.
;   TITLE OF INVENTION: Assays for Protein Kinases Using
;   TITLE OF INVENTION: Fluorescent Protein Substrates
;   NUMBER OF SEQUENCES: 48
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Townsend and Townsend and Crew LLP
;     STREET: Two Embarcadero Center, Eighth Floor
;     CITY: San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94111-3834
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/263,975
;     FILING DATE:
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/679,865
;     FILING DATE: 16-JUL-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Storella, John S.
;     REGISTRATION NUMBER: 32,944
;     REFERENCE/DOCKET NUMBER: 023072-069000
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (415) 576-0200
;     TELEFAX: (415) 576-0300
;   INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 238 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-09-263-975-2

Query Match             98.6%;  Score 1258;  DB 3;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 8.4e-129;
Matches  235;           Conservative    0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 9
US-08-727-452-2
;   Sequence 2, Application US/08727452A
;   Patent No. 6319669
;   GENERAL INFORMATION:
;   APPLICANT: Tsien, Roger Y.
;   APPLICANT: Heim, Roger
;   TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR DETECTION OF ANALYTES
;   FILE REFERENCE: 07257/032001
;   CURRENT APPLICATION NUMBER: US/08/727,452A
;   CURRENT FILING DATE: 1996-03-20
;   EARLIER APPLICATION NUMBER: PCT/US95/14692
;   EARLIER FILING DATE: 1995-11-13
;   EARLIER APPLICATION NUMBER: US 07/337,915
;   EARLIER FILING DATE: 1994-11-10
;   NUMBER OF SEQ ID NOS: 5
;   SOFTWARE: FastSEQ for Windows Version 3.0
;   SEQ ID NO 2
;     LENGTH: 238
;     TYPE: PRT
;     ORGANISM: Aequorea victoria
US-08-727-452-2

Query Match             98.6%;  Score 1258;  DB 3;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 8.4e-129;
Matches  235;           Conservative    0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

```
RESULT 10
US-09-418-785-1
; Sequence 1, Application US/09418785
; Patent No. 6414119
; GENERAL INFORMATION:
; APPLICANT: Fisher, Hugh
; TITLE OF INVENTION: Rapidly Greening, Low Oxygen Mutant of
; TITLE OF INVENTION: the Aequorea victoria Green Fluorescent Protein
; FILE REFERENCE: RuCC 99-0011
; CURRENT APPLICATION NUMBER: US/09/418,785
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/104,563
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
; PUBLICATION INFORMATION:
; AUTHORS: Prasher, D.C. et al.
; TITLE: Primary structure of the Aequorea victorea green-f
; JOURNAL: Gene
; VOLUME: 111
; PAGES: 229-233
; DATE: 1992-01-01
; DATABASE ACCESSION NUMBER: Genbank M62653
; DATABASE ENTRY DATE: 1993-04-26
US-09-418-785-1

Query Match             98.6%;  Score 1258;  DB 4;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 8.4e-129;
Matches   235;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              |||.||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||||||||||||||||||.||||||||||||||||
Db        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 11
US-09-129-192C-2
; Sequence 2, Application US/09129192C
; Patent No. 6495664
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AURO1270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/09/129,192C
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea
US-09-129-192C-2

Query Match             98.6%;  Score 1258;  DB 4;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 8.4e-129;
Matches   235;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              |||.||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||||||||||||||||||.||||||||||||||||
Db        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 12
US-09-129-192C-74
; Sequence 74, Application US/09129192C
; Patent No. 6495664
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AURO1270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/09/129,192C
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SEQ ID NO 74
; LENGTH: 238
; TYPE: PRT
; SOFTWARE: PatentIn version 3.0
; ORGANISM: Aequorea green fluorescent protein phosphorylation mutant
US-09-129-192C-74

Query Match             98.6%;  Score 1258;  DB 4;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 8.4e-129;
Matches   235;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              |||.||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||||||||||||||||||.||||||||||||||||
Db        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 13
US-09-602-641-2
; Sequence 2, Application US/09602641
; Patent No. 6608189
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/602,641
```

```
;  CURRENT FILING DATE: 2000-06-22
;  PRIOR APPLICATION NUMBER: 09/172,063
;  PRIOR FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 38
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 2
;  LENGTH: 238
;  TYPE: PRT
;  ORGANISM: Aequorea victoria
US-09-602-641-2

Query Match    98.6%;  Score 1258;  DB 4;  Length 238;
Best Local Similarity  98.7%;  Pred. No. 8.4e-129;
Matches  235;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy  61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
       ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 14
US-09-704-463-2
;  Sequence 2, Application US/09704463
;  Patent No. 6627449
;  GENERAL INFORMATION:
;  APPLICANT: Tsien, Roger Y.
;  APPLICANT: Miyawaki, Atsushi
;  APPLICANT: Llopis, Juan
;  TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
;  TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
;  FILE REFERENCE: 07257/067001
;  CURRENT APPLICATION NUMBER: US/09/704,463
;  CURRENT FILING DATE: 2000-10-31
;  PRIOR APPLICATION NUMBER: 09/094,359
;  PRIOR FILING DATE: 1998-06-09
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 2
;  LENGTH: 238
;  TYPE: PRT
;  ORGANISM: Aequorea victoria
US-09-704-463-2

Query Match    98.6%;  Score 1258;  DB 4;  Length 238;
Best Local Similarity  98.7%;  Pred. No. 8.4e-129;
Matches  235;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy  61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
       ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 15
US-10-024-686A-2
;  Sequence 2, Application US/10024686A
;  Patent No. 6800733
;  GENERAL INFORMATION:
;  APPLICANT: Tsien, Roger Y.
;  APPLICANT: Heim, Roger
;  TITLE OF INVENTION: Modified Green Fluorescent Proteins
;  FILE REFERENCE: 3754/0861
;  CURRENT APPLICATION NUMBER: US/10/024,686A
;  CURRENT FILING DATE: 2001-12-17
;  PRIOR APPLICATION NUMBER: 09/057,995
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 08/753,144
;  PRIOR FILING DATE: 1996-11-20
;  PRIOR APPLICATION NUMBER: 08/727,452
;  PRIOR FILING DATE: 1997-03-20
;  PRIOR APPLICATION NUMBER: PCT/US95/14692
;  PRIOR FILING DATE: 1995-11-13
;  PRIOR APPLICATION NUMBER: 08/337,915
;  PRIOR FILING DATE: 1994-11-10
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 2
;  LENGTH: 238
;  TYPE: PRT
;  ORGANISM: Aequorea victoria
US-10-024-686A-2

Query Match    98.6%;  Score 1258;  DB 4;  Length 238;
Best Local Similarity  98.7%;  Pred. No. 8.4e-129;
Matches  235;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy  61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
       ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

Search completed: September 16, 2005, 17:22:10
Job time : 33 secs
```

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:        September 16, 2005, 17:20:18 ; Search time 113 Seconds
                                             (without alignments)
                                   852.856 Million cell updates/sec

Title:          US-10-757-624-3
Perfect score:  1276
Sequence:       1 MSKGEELFTGVVPILVELDG..........VLLGFVTAAGITHGMDELYK 238

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters:      1812044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications_AA:*
         1:  /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
         2:  /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
         3:  /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
         4:  /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
         5:  /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
         6:  /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
         7:  /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
         8:  /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
         9:  /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
        10:  /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
        11:  /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
        12:  /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
        13:  /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
        14:  /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
        15:  /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
        16:  /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
        17:  /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
        18:  /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
        19:  /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
        20:  /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
        21:  /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
        22:  /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

                               SUMMARIES
```

Result No.	Score	Query Match %	Match Length	DB	ID	Description
1	1276	100.0	238	10	US-09-967-301-3	Sequence 3, Appli
2	1276	100.0	238	10	US-10-757-624-3	Sequence 3, Appli
3	1265	99.1	238	10	US-09-967-301-4	Sequence 4, Appli
4	1265	99.1	238	9	US-10-757-624-4	Sequence 4, Appli
5	1258	98.6	238	9	US-09-884-681-2	Sequence 2, Appli
6	1258	98.6	238	10	US-09-967-301-2	Sequence 2, Appli
7	1258	98.6	238	13	US-10-024-686-2	Sequence 2, Appli
8	1258	98.6	238	13	US-10-057-505-2	Sequence 2, Appli
9	1258	98.6	238	14	US-10-293-580-2	Sequence 2, Appli
10	1258	98.6	238	14	US-10-293-580-74	Sequence 74, Appl
11	1258	98.6	238	15	US-10-457-982-2	Sequence 2, Appli
12	1258	98.6	238	16	US-10-724-178-2	Sequence 2, Appl
13	1258	98.6	238	16	US-10-757-624-2	Sequence 2, Appli
14	1258	98.6	238	17	US-10-505-486-3	Sequence 34, App
15	1258	98.6	1125	16	US-10-845-936A-34	Sequence 2, Appli
16	1255	98.4	238	16	US-09-899-954B-2	Sequence 4, Appl
17	1254	98.3	238	9	US-09-920-922-4	Sequence 1, Appl
18	1254	98.3	238	9	US-09-852-000-1	Sequence 125, App
19	1254	98.3	238	10	US-09-900-345A-125	Sequence 2, Appl
20	1254	98.3	238	10	US-09-866-538-2	Sequence 2, Appl
21	1254	98.3	238	10	US-09-794-308-2	Sequence 2, Appl
22	1254	98.3	238	10	US-09-865-291-2	Sequence 10, Appl
23	1254	98.3	238	14	US-10-121-258-10	Sequence 6, Appl
24	1254	98.3	238	14	US-10-221-461-6	Sequence 15, Appl
25	1254	98.3	238	14	US-10-321-195-15	Sequence 10, Appl
26	1254	98.3	238	14	US-10-305-765-10	Sequence 159, App
27	1254	98.3	238	14	US-10-305-765-159	Sequence 10, Appl
28	1254	98.3	238	14	US-10-305-633-10	Sequence 159, App
29	1254	98.3	238	14	US-10-305-633-159	Sequence 8, Appl
30	1254	98.3	238	15	US-10-132-067-8	Sequence 1, Appl
31	1254	98.3	238	15	US-10-370-570-1	Sequence 53, Appl
32	1254	98.3	238	15	US-10-370-570-53	Sequence 42, Appl
33	1254	98.3	238	15	US-10-423-688A-42	Sequence 4, Appl
34	1254	98.3	238	15	US-10-668-168-4	Sequence 3, Appl
35	1254	98.3	238	16	US-10-333-680-3	Sequence 1, Appl
36	1254	98.3	238	16	US-10-753-405-1	Sequence 35, Appl
37	1254	98.3	238	16	US-10-699-113-35	Sequence 2, Appl
38	1254	98.3	238	16	US-10-845-484-2	Sequence 2, Appl
39	1254	98.3	238	16	US-10-885-988-2	Sequence 2, Appl
40	1254	98.3	238	17	US-10-857-622-2	Sequence 10, Appl
41	1254	98.3	238	18	US-10-931-304-10	Sequence 125, App
42	1254	98.3	238	20	US-11-041-806-125	Sequence 6, Appl
43	1253	98.2	238	15	US-10-668-168-6	Sequence 35, Appl
44	1253	98.2	238	16	US-10-676-428-35	Sequence 8, Appl
45	1252	98.1	244	9	US-09-899-954B-8	

```
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.


                              ALIGNMENTS

RESULT 1
US-09-967-301-3
; Sequence 3, Application US/09967301
; Publication No. US2003017585A1
; GENERAL INFORMATION:
;  APPLICANT: Stubbs, Simon L.
;  APPLICANT: Jones, Anne E.
;  APPLICANT: Michael, Nigel P.
;  APPLICANT: Thomas, Nicholas
;  TITLE OF INVENTION: Fluorescent Proteins
;  FILE REFERENCE: PA0111
;  CURRENT APPLICATION NUMBER: US/09/967,301
;  CURRENT FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: GB 0109858.1
;  PRIOR FILING DATE: 2001-09-28
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 3
;  LENGTH: 238
;  TYPE: PRT
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: synthetic
;  OTHER INFORMATION: protein
US-09-967-301-3

Query Match             100.0%;  Score 1276;  DB 10;  Length 238;
Best Local Similarity   100.0%;  Pred. No. 1.3e-117;
Matches  238;  Mismatches   0;  Conservative   0;  Indels   0;  Gaps   0;

Qy        1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
```

```
Qy   61  VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
```

RESULT 2
US-10-757-624-3

; Sequence 3, Application US/10757624
; Publication No. US2004138420A1
; GENERAL INFORMATION:
; APPLICANT: Stubbs, Simon L. J.
; APPLICANT: Jones, Anne E.
; APPLICANT: Michael, Nigel P.
; APPLICANT: Thomas, Nicholas
; TITLE OF INVENTION: Fluorescent Proteins
; FILE REFERENCE: PA0111
; CURRENT APPLICATION NUMBER: US/10/757,624
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 09/967,301
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: GB 0109858.1
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.2
;
; SEQ ID NO 3
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic protein
US-10-757-624-3

```
Query Match     100.0%;  Score 1276;  DB 16;  Length 238;
Best Local Similarity   100.0%;  Pred. No. 1.3e-117;
Matches  238;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61  VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
```

RESULT 3
US-09-967-301-4

; Sequence 4, Application US/09967301
; Publication No. US20030175859A1
; GENERAL INFORMATION:
; APPLICANT: Stubbs, Simon L.
; APPLICANT: Jones, Anne E.
; APPLICANT: Michael, Nigel P.
; APPLICANT: Thomas, Nicholas
; TITLE OF INVENTION: Fluorescent Proteins
; FILE REFERENCE: PA0111
; CURRENT APPLICATION NUMBER: US/09/967,301

; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: GB 0109858.1
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 4
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: protein
US-09-967-301-4

```
Query Match      99.1%;  Score 1265;  DB 10;  Length 238;
Best Local Similarity    99.2%;  Pred. No. 1.6e-116;
Matches  236;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61  VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 4
US-10-757-624-4

; Sequence 4, Application US/10757624
; Publication No. US2004138420A1
; GENERAL INFORMATION:
; APPLICANT: Stubbs, Simon L. J.
; APPLICANT: Jones, Anne E.
; APPLICANT: Michael, Nigel P.
; APPLICANT: Thomas, Nicholas
; TITLE OF INVENTION: Fluorescent Proteins
; FILE REFERENCE: PA0111
; CURRENT APPLICATION NUMBER: US/10/757,624
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 09/967,301
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: GB 0109858.1
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.2
;
; SEQ ID NO 4
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic protein
US-10-757-624-4

```
Query Match      99.1%;  Score 1265;  DB 16;  Length 238;
Best Local Similarity    99.2%;  Pred. No. 1.6e-116;
Matches  236;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLITLKFICTTGKLPVPWPTL 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61  VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
```

```
Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 5
US-09-884-681-2
; Sequence 2, Application US/09884681
; Patent No. US2002061546A1
; GENERAL INFORMATION:
;   APPLICANT: Tsien, Roger Y.
;   Cubitt, Andrew B.
;   TITLE OF INVENTION: Assays for Protein Kinases Using
;   Fluorescent Protein Substrates
;   NUMBER OF SEQUENCES: 48
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Townsend and Townsend and Crew LLP
;   STREET: Two Embarcadero Center, Eighth Floor
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94111-3834
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/884,681
;   FILING DATE: 19-Jun-2001
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/679,865
;   FILING DATE: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Storella, John S.
;   REGISTRATION NUMBER: 32,944
;   REFERENCE/DOCKET NUMBER: 02307Z-069000
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 576-0200
;   TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 238 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-884-681-2

```
Query Match      98.6%;   Score 1258;  DB 9;  Length 238;
Best Local Similarity   98.7%;   Pred. No. 7.9e-116;
Matches  235;   Conservative   0;   Mismatches   3;   Indels  0;   Gaps  0;

Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
```

```
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 6
US-09-967-301-2
; Sequence 2, Application US/09967301
; Publication No. US20030175859A1
; GENERAL INFORMATION:
;   APPLICANT: Stubbs, Simon L.
;   APPLICANT: Jones, Anne E.
;   APPLICANT: Michael, Nigel P.
;   APPLICANT: Thomas, Nicholas
;   TITLE OF INVENTION: Fluorescent Proteins
;   FILE REFERENCE: PA0111
;   CURRENT APPLICATION NUMBER: US/09/967,301
;   CURRENT FILING DATE: 2001-09-28
;   PRIOR APPLICATION NUMBER: GB 0109858.1
;   PRIOR FILING DATE: 2001-04-23
;   NUMBER OF SEQ ID NOS: 19
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Aequorea victoria
US-09-967-301-2

```
Query Match      98.6%;   Score 1258;  DB 10;  Length 238;
Best Local Similarity   98.7%;   Pred. No. 7.9e-116;
Matches  235;   Conservative   0;   Mismatches   3;   Indels  0;   Gaps  0;

Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 7
US-10-024-686-2
; Sequence 2, Application US/10024686
; Publication No. US20020123113A1
; GENERAL INFORMATION:
;   APPLICANT: Tsien, Roger Y.
;   Heim, Roger
;   TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Fish & Richardson P.C.
;   STREET: 4225 Executive Square, Suite 1400
;   CITY: La Jolla
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 92037
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: Windiws95
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/024,686
;   FILING DATE: 17-Dec-2001
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 09/057,995

```
 ;   FILING DATE: <Unknown>
 ;   APPLICATION NUMBER: 08/727,452
 ;   FILING DATE: 10-OCT-1996
 ;   APPLICATION NUMBER: US95/14692
 ;   FILING DATE: 13-NOV-1995
 ;   APPLICATION NUMBER: 08/337,915
 ;   FILING DATE: 10-NOV-1994
 ;   ATTORNEY/AGENT INFORMATION:
 ;     NAME: Haile, Lisa A.
 ;     REGISTRATION NUMBER: 38,347
 ;     REFERENCE/DOCKET NUMBER: 07257/032002
 ;   TELECOMMUNICATION INFORMATION:
 ;     TELEPHONE: 619/678-5070
 ;     TELEFAX: 619/678-5099
 ; INFORMATION FOR SEQ ID NO: 2:
 ;   SEQUENCE CHARACTERISTICS:
 ;     LENGTH: 238 amino acids
 ;     TYPE: amino acid
 ;     TOPOLOGY: linear
 ;   MOLECULE TYPE: protein
 ;   FRAGMENT TYPE: internal
 ;   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-024-686-2

Query Match             98.6%;  Score 1258; DB 13; Length 238;
Best Local Similarity   98.7%;  Pred. No. 7.9e-116;
Matches   235;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy        121 NRIELKGIDFKEDGNIIGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 8
US-10-057-505-2
; Sequence 2, Application US/10057505
; Publication No. US20020164574A1
; GENERAL INFORMATION:
;   APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
;   APPLICANT: AURORA BIOSCIENCES CORPORATION
;   APPLICANT: TSIEN, Roger
;   APPLICANT: HEIM, Roger
;   APPLICANT: CUBITT, Andrew
;   TITLE OF INVENTION: TANDEM FLUORESCENT PROTEIN CONSTRUCTS
;   FILE REFERENCE: REGEN1260-3
;   CURRENT APPLICATION NUMBER: US/10/057,505
;   CURRENT FILING DATE: 2002-01-25
;   PRIOR APPLICATION NUMBER: US 08/792,553
;   PRIOR FILING DATE: 1997-01-31
;   PRIOR APPLICATION NUMBER: US 09/396,003
;   PRIOR FILING DATE: 1999-09-13
;   NUMBER OF SEQ ID NOS: 31
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Aequorea victoria
US-10-057-505-2

Query Match             98.6%;  Score 1258; DB 13; Length 238;
Best Local Similarity   98.7%;  Pred. No. 7.9e-116;
```

```
Matches   235;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy        121 NRIELKGIDFKEDGNIIGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 9
US-10-293-580-2
; Sequence 2, Application US/10293580
; Publication No. US20030170767A1
; GENERAL INFORMATION:
;   APPLICANT: Aurora Biosciences Corporation
;   APPLICANT: Cubitt, Andrew B.
;   TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
;   FILE REFERENCE: AURO1270 (08366/031001)
;   CURRENT APPLICATION NUMBER: US/10/293,580
;   CURRENT FILING DATE: 2002-11-12
;   PRIOR APPLICATION NUMBER: US/09/129,192
;   PRIOR FILING DATE: 1998-07-24
;   NUMBER OF SEQ ID NOS: 74
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Aequorea
US-10-293-580-2

Query Match             98.6%;  Score 1258; DB 14; Length 238;
Best Local Similarity   98.7%;  Pred. No. 7.9e-116;
Matches   235;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy         61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy        121 NRIELKGIDFKEDGNIIGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 10
US-10-293-580-74
; Sequence 74, Application US/10293580
; Publication No. US20030170767A1
; GENERAL INFORMATION:
;   APPLICANT: Aurora Biosciences Corporation
;   APPLICANT: Cubitt, Andrew B.
;   TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
;   FILE REFERENCE: AURO1270 (08366/031001)
;   CURRENT APPLICATION NUMBER: US/10/293,580
;   CURRENT FILING DATE: 2002-11-12
;   PRIOR APPLICATION NUMBER: US/09/129,192
```

; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 74
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea green fluorescent protein phosphorylation mutant
US-10-293-580-74

Query Match   98.6%;   Score 1258; DB 14;   Length 238;
Best Local Similarity 98.7%;   Pred. No. 7.9e-116;
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy 61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
       |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
       |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 11
US-10-457-982-2
; Sequence 2, Application US/10457982
; Publication No. US20030212265A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/10/457,982
; CURRENT FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: US/09/602,641
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/172,063
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-10-457-982-2

Query Match   98.6%;   Score 1258; DB 15;   Length 238;
Best Local Similarity 98.7%;   Pred. No. 7.9e-116;
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy 61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
       |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
       |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 12
US-10-724-178-2
; Sequence 2, Application US/10724178
; Publication No. US20040137528A1
; GENERAL INFORMATION:
; APPLICANT: Odyssey Thera, Inc.
; APPLICANT: Michnick, Stephen
; APPLICANT: MacDonald, Marnie
; APPLICANT: Lamerdin, Jane
; TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT
; TITLE OF INVENTION: COMPLEMENTATION ASSAYS
; FILE REFERENCE: ODDY007
; CURRENT APPLICATION NUMBER: US/10/724,178
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US 60/461,133
; PRIOR FILING DATE: 2003-04-09
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-10-724-178-2

Query Match   98.6%;   Score 1258; DB 16;   Length 238;
Best Local Similarity 98.7%;   Pred. No. 7.9e-116;
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy 61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
       |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
       |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 13
US-10-757-624-2
; Sequence 2, Application US/10757624
; Publication No. US20040138420A1
; GENERAL INFORMATION:
; APPLICANT: Stubbs, Simon L. J.
; APPLICANT: Jones, Anne B.
; APPLICANT: Michael, Nigel P.
; APPLICANT: Thomas, Nicholas
; TITLE OF INVENTION: Fluorescent Proteins
; FILE REFERENCE: PA0111
; CURRENT APPLICATION NUMBER: US/10/757,624
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 09/967,301
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: GB 0109858.1
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria

US-10-757-624-2

Query Match   98.6%;   Score 1258; DB 16;   Length 238;
Best Local Similarity 98.7%;   Pred. No. 7.9e-116;
Matches 235; Conservative   0; Mismatches   3;   Indels   0;   Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        ||||.|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||.||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 14
US-10-505-486-3
; Sequence 3, Application US/10505486
; Publication No. US2005011863 9A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Determination of a ligand
; FILE REFERENCE: P03-0006PCT
; CURRENT APPLICATION NUMBER: US/10/505,486
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: JP 2002-45728
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: JP 2002-213949
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: JP 2002-298237
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 233
; SEQ ID NO 3
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea coerulescens
US-10-505-486-3

Query Match   98.6%;   Score 1258; DB 17;   Length 238;
Best Local Similarity 98.7%;   Pred. No. 7.9e-116;
Matches 235; Conservative   0; Mismatches   3;   Indels   0;   Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        ||||.|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||.||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 15
US-10-845-936A-34
; Sequence 34, Application US/10845936A
; Publication No. US20040234609A1
; GENERAL INFORMATION:
; APPLICANT: Collier, Katherine D.

; APPLICANT: Cuevas, William A.
; APPLICANT: Kumar, Manoj A.
; TITLE OF INVENTION: Repeat Sequence Protein Polymer Active Agent Conjugates, Methods
; TITLE OF INVENTION: and Uses
; FILE REFERENCE: DOC0068PA/DC5058/GC792
; CURRENT APPLICATION NUMBER: US/10/845,936A
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: 60/470,464
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 1125
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: GFP-SELP47K: silk, elastin and green fluorescent protein peptides
US-10-845-936A-34

Query Match   98.6%;   Score 1258; DB 16;   Length 1125;
Best Local Similarity 98.7%;   Pred. No. 6.7e-115;
Matches 235; Conservative   0; Mismatches   3;   Indels   0;   Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 63

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        ||||.|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 123

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 183

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||.||||||||||||||||
Db 184 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 241

Search completed: September 16, 2005, 17:39:31
Job time : 114 secs

OM protein - protein search, using sw model

```
              GenCore version 5.1.6
    Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:   September 16, 2005, 17:06:47 ; Search time 24.5 Seconds
                  (without alignments)
                  934.677 Million cell updates/sec

Title:     US-10-757-624-3
Perfect score: 1276
Sequence:  1 MSKGEELFTGVVPILVELDG.........VLLGFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62
           Gapop 10.0 , Gapext 0.5

Database : PIR 79:*
           1: pir1:*
           2: pir2:*
           3: pir3:*
           4: pir4:*

Searched:   283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:     283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1248	97.8	238	1	JQ1514	green-fluorescent
2	107	8.4	785	2	H72228	hypothetical prote
3	99.5	7.8	887	2	B82590	leucyl-tRNA synthe
4	92.5	7.2	534	1	NICLMA	nitrogenase (EC 1.
5	90	7.1	1224	1	ERHUAH	coatomer complex a
6	89.5	7.0	861	2	H64102	leucine-tRNA ligas
7	88.5	6.9	370	2	E70390	iron-sulfur cofact
8	88	6.9	353	2	E84941	imidazoleglycerol-
9	87.5	6.9	797	2	JC4078	protective surface
10	87.5	6.9	808	2	F64102	protective surface
11	87.5	6.9	2222	1	A36028	DNA-directed DNA p
12	87	6.8	357	2	G81355	tRNA (uracil-5-)-m
13	87	6.8	439	2	JH0414	synaptogamin o-p65
14	86.5	6.8	263	2	S53488	water-stress-induc
15	86.5	6.8	860	2	AC0582	leucyl-tRNA synthe
16	86.5	6.8	2573	2	D71614	hypothetical prote
17	85.5	6.7	655	2	D83917	DNA topoisomerase
18	85	6.7	632	2	T06586	DNA-binding protei
19	84.5	6.6	700	1	CIHUH2	calpain (EC 3.4.22
20	84.5	6.6	889	2	JC5576	inter-alpha-trypsi
21	84	6.6	281	2	AD2052	hypothetical prote
22	84	6.6	531	2	C95338	hypothetical prote
23	84	6.6	874	2	JC4930	S-layer protein pr
24	83.5	6.5	613	2	A99552	oligoendopeptidase
25	83	6.5	578	1	I40794	dihydrolipoamide d
26	82.5	6.5	312	2	C81710	thioredoxin reduct
27	82.5	6.5	422	2	B24815	calpain (EC 3.4.22
28	82.5	6.5	502	2	B90103	T-complex protein
29	82.5	6.5	941	2	S29043	cellulase (EC 3.2.
30	82.5	6.5	1134	2	A60234	IgA Fc receptor pr
31	82.5	6.5	1164	1	FCSOAG	IgA Fc receptor pr
32	82	6.4	470	2	T31049	hypothetical prote
33	82	6.4	774	2	T39539	alpha-amylase homo
34	82	6.4	865	2	T11852	lipoxygenase (EC 1
35	81.5	6.4	310	2	S68225	synergohymenotropi
36	81.5	6.4	425	2	C97354	hypothetical prote
37	81.5	6.4	564	2	B81317	ABC-type transport
38	81.5	6.4	788	1	JDVLHH	DNA-directed DNA p
39	81	6.3	336	2	C64468	hypothetical prote
40	81	6.3	351	2	A88103	protein W10G11.7 [
41	81	6.3	461	2	T06936	photosystem II chl
42	81	6.3	1516	2	T41235	probable myosin he
43	80.5	6.3	389	2	G84245	NADH dehydrogenase
44	80.5	6.3	393	2	C64613	conserved hypothet
45	80.5	6.3	653	1	AI1535	fructose-bisphosph

ALIGNMENTS

```
RESULT 1
JQ1514
green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)
C;Species: Aequorea victoria
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: JS0692; JQ1514; PQ0335; S4869; S51330; S51331
R;Prasher, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.
Gene 111, 229-233, 1992
A;Title: primary structure of the Aequorea victoria green-fluorescent protein.
A;Reference number: JQ1514; MUID:92175527; PMID:1347277
A;Accession: JS0692
A;Molecule type: DNA
A;Residues: 1-107,'S',109-238 <PRA1>
A;Cross-reference: UNIPROT:P42212; UNIPROT:Q17106; UNIPROT:Q17105; GB:M62654; NID:g1556
A;Accession: JQ1514
A;Molecule type: mRNA
A;Residues: 1-99,'F',101-140,'L',142-218,'V',220-238 <PRA2>
A;Cross-references: GB:M62653; NID:g155660; PIDN:AAA27721.1; PID:g155661
A;Accession: PQ0335
A;Molecule type: protein
A;Residues: 46-64;74-122;132-151;154-183;185-200 <PRA3>
R;Inouye, S.; Tsuji, F.I.
FEBS Lett. 351, 211-214, 1994
A;Title: Evidence for redox forms of the Aequorea green fluorescent protein.
A;Reference number: S48693; MUID:94364470; PMID:8082767
A;Accession: S48693
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-24,'Q',26-156,'P',158-171,'K',173-238 <INO>
A;Cross-references: GB:L29345; NID:g606383; PIDN:AAA58246.1; PID:g606384
R;Watkins, J.N.; Campbell, A.K.
submitted to the EMBL Data Library, January 1995
A;Reference number: S51330
A;Accession: S51330
A;Molecule type: mRNA
A;Residues: 1-13,'V',15-24,'Q',26-44,'N',46-153,'G',155-156,'P',158-171,'K',173-227,'R',
A;Cross-references: EMBL:X83959; NID:g634008; PIDN:CAA58789.1; PID:g634009
A;Experimental source: clone gfp1
A;Accession: S51331
A;Molecule type: mRNA
A;Residues: 1-24,'Q',26-29,'R',31-83,'L',85-153,'G',155-156,'P',158-171,'K',173-208,'Q',
A;Cross-references: EMBL:X83960; NID:g634010; PIDN:CAA58790.1; PID:g634011
A;Experimental source: clone gfp2
R;Yang, F.; Moss, L.G.; Phillips Jr., G.N.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65692; PDB:IGFL
A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 'A',2-79,'R',81-5
A;Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli
R;Yang, F.; Moss, L.G.; Phillips Jr., G.N.
Nat. Biotechnol. 14, 1246-1251, 1996
A;Title: The molecular structure of green fluorescent protein.
A;Reference number: A58953; MUID:98294543; PMID:9631087
```

A;Contents: annotation; X-ray crystallography, 1.9 angstroms
C;Comment: This protein is excited by the photoprotein aequorin (see PIR:AQJFNV) emittir
C;Comment: The chromophore of this protein is formed by modification of Ser-dehydro-Tyr-
A;Genetics:
A;Gene: GFP
A;Introns: 69/3; 167/3
C;Superfamily: green-fluorescent protein
C;Keywords: chromoprotein; luminescence
F;65-67/Cross-link: 5-imidazolinone (Ser-Gly) #status experimental
F;66/Modified site: dehydrotyrosine (Tyr) #status experimental

Query Match   97.8%;   Score 1248; DB 1;   Length 238;
Best Local Similarity 97.1%;   Pred. No. 5.3e-99;
Matches 231; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFYKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKMEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMILLEFVTAAGITHGMDELYK 238

RESULT 2
H72228
hypothetical protein TM1624 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: H72228
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-785 <ARN>
A;Cross-references: UNIPROT:Q9X1V9; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AAD3669
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1624

Query Match   8.4%;   Score 107;   DB 2;   Length 785;
Best Local Similarity 19.7%;   Pred. No. 0.45;
Matches 46; Conservative 33; Mismatches 70; Indels 84; Gaps 7;

Qy   2 SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLV 61
           .: |||: || :||:| || :
Db   15 NEGRFSFEGTVPGVVQAD---------------------------LVRKGLLPHYVGM- 46

Qy   62 TTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVN 121
           ||| | || | | | |||||||||||||||||||||||||
Db   47 ----------NEDLFKEIEDREWIYEREFEFKEDVKEGERVDLVFEGVDTLS 88

Qy   122 RIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLADH 181
           |||||||||||||||||||:||:||:|| |||||||||||||||||||||
Db   89 DVYLNGVYL---GSTEDMFIEYRFDVTNVL----KEKNHLKVYIK------SPIRVPKT 134

Qy   182 YQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMD 234
           :|||||||||||||||||||||||||||||||||||||||||||||||:||
Db   135 LEQNYGVLGGP-----------EDP-------IRGYIRKAQYSYGWD 163

RESULT 3
E82590
leucyl-tRNA synthetase XF2176 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: E82590
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title:The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: E82590
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-887 <SIM>
A;Cross-references: GB:AE004031; GB:AE003849; NID:g9107309; PIDN:AAF84975.1; GSPDB:GN001:
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Za
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2176
C;Superfamily: leucine-tRNA ligase

Query Match   7.8%;   Score 99.5; DB 2;   Length 887;
Best Local Similarity 23.2%;   Pred. No. 2.3;
Matches 46; Conservative 28; Mismatches 67; Indels 57; Gaps 10;

Qy   49 TTGKLPVPWPTLVTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNY-- 106
           :||| | :||| :| ||| | :   ::
Db   329 TNEQLPV-WVANFVLMAYGTGAVMAVPGHDQRDQEF--ANKYGLPIRQVIALKBPKNQDE 385

Qy   107 ------------------KTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYI 152
                             :||: |:||:: |||
Db   386 STWEPDVWRDWYADKTR---EFE---LINSAEFDGLDVQDAFEVLAERFE---------- 429

Qy   153 MADKQKNG-IKVNFKIRHNIEDGGVQLADHYQQNTPI-----------------GDGPVLLPDN 198
           : :: |: ||:|: | :: ||:||| :|
Db   430 ---RQGRGQRRVNYRLR-----DWGVSRQRYWGCPIPVIYCPTCGAVPVPEDQLPVILPEN 482

Qy   199 -HYLSTQSALSKDPNEKR 215
           :|: | ::
Db   483 VAFSGTGSPIKTDPEWRK 500

RESULT 4
NICLMA
nitrogenase (EC 1.18.6.1) molybdenum-iron protein alpha chain [validated] - Clostridium p
N;Alternate names: dinitrogenase alpha chain; nitrogenase component I alpha chain
C;Species: Clostridium pasteurianum
C;Date: 01-Sep-1981 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C;Accession: S07389; I40816; A00545
R;Wang, S.Z.; Chen, J.S.; Johnson, J.L.
Nucleic Acids Res. 15, 3935, 1987
A;Title: Nucleotide and deduced amino acid sequences of nifD encoding the alpha-subunit c
A;Reference number: S07389; MUID:87231095; PMID:3473447
A;Accession: S07389
A;Molecule type: DNA
A;Residues: 1-534 <WAN>
A;Cross-references: UNIPROT:P00467; EMBL:Y00155; NID:g40583; PIDN:CAA68349.1; PID:g58099£
R;Chen, K.C.

J. Bacteriol. 166, 162-172, 1986
A;Title: Structural features of multiple nifH-like sequences and very biased codon usage
A;Reference number: I40814; MUID:86168010; PMID:3457003
A;Accession: I40816
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-167 <CH3>
A;Cross-references: GB:M21537; NID:g144870; PIDN:AAA83531.1; PID:g551775
R;Hase, T.; Nakano, T.; Matsubara, H.; Zumft, W.G.
J. Biochem. 90, 295-298, 1981
A;Title: Correspondence of the larger subunit of the MoFe-protein in clostridial nitroge
A;Reference number: A00545; MUID:82030699; PMID:7026551
A;Accession: A00545
A;Molecule type: protein
A;Residues: 2-41,'K',43-94,'D',96-180 <HAS>
R;Kim, J.; Woo, D.; Rees, D.C.
submitted to the Brookhaven Protein Data Bank, March 1993
A;Reference number: A51301; PDB:1MIO
A;Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 2-425,'A',427-526
C;Comment: The key enzymatic reactions in nitrogen fixation are catalyzed by the nitroge
and the iron protein (also called component II or nitrogenase reductase).
C;Comment: Each alpha/beta dimer covalently binds one molybdenum-iron-sulfur cluster and
C;Genetics:
A;Gene: nifD
A;Start codon: GTG
C;Complex: heterotetramer of two alpha and two beta chains
C;Function:
A;Description: the enzyme complex catalyzes the reduction of dinitrogen to 2 molecules o
A;Pathway: nitrogen fixation
A;Superfamily: dinitrogenase alpha chain; nitrogenase vanadium-iron protein alpha chain
C;Superfamily:
C;Keywords: 4Fe-4S; ATP; heterotetramer; iron-sulfur protein; metalloprotein; molybdenum
F;2-534/Product: nitrogenase molybdenum-iron protein alpha chain #status experimental <M
F;5-522/Domain: nitrogenase vanadium-iron protein alpha chain homology <VIA>
F;53,79,145/Binding site: 4Fe-4S cluster 1 (Cys) (covalent) #status experimental
F;79/Binding site: 4Fe-4S cluster 2 (Cys) (covalent) #status experimental
F;262/Binding site: homocitryl Mo-7Fe-8S cluster (Cys) (covalent) #status experimental
F;482/Binding site: homocitryl Mo-7Fe-8S cluster molybdenum (His) (ligand) #status exper

Query Match   7.2%;   Score 92.5;   DB 1;   Length 534;
Best Local Similarity   19.9%;   Pred. No. 4.7;
Matches   48;   Conservative   39;   Mismatches   97;   Indels   57;   Gaps   8;

```
Qy   17 ELDGDVNGHKFSVSGEGEGDATYGKL------------TLKPIC-TTGKLPVPW-- 57
           |:|  ||||||:                :||
Db  224 EMDRVLEKIGYHVNATLTGDATYEKVQNADKADLNLVQCHRSINYIAEMETKYGIPWIK 283

Qy   58 ------PTLVTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFPKDDGNY-KTRA 110
                  |  |:|         |:
Db  284 CNFIGVDGIVETLRDMAKCFDD-PELTKR--------TEVIABBIAAIQDDLDYFKEKL 334

Qy  111 EVK--------FEGDTLVNRIELKGIDFKEDGNILGHKLEY-----------NYNSH 148
            |:|        ||||||||||||||||||||||||
Db  335 QGKTACLYVGGSRSHTYMNMLKSFGVDSLVAGFBPAHRDDYEGREVIPTIKIDADSKNIP 394

Qy  149 NVYIMADKQKNGIKVNFKIRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS 208
           |||||||||||||||||||:|    |||||||||||  :|   ||||||||
Db  395 EITVTPDEQKYRVVIPEDKVEBLKKAGVPLISSYGGMKEMHDGTILIDDMNHHDMEVVLE 454

Qy  209 K 209
Db  455 K 455
```

RESULT 5
ERHUAH
coatomer complex alpha chain homolog - human
N;Alternate names: HEP-COP; xenopsin homolog; xenopsin-related peptide precursor
N;Contains: xenin 25; xenopsin-related peptide
.C;Species: Homo sapiens (man)
C;Date: 10-May-1996 #sequence_revision 08-Nov-1996 #text_change 09-Jul-2004
C;Accession: JC4668; A44317
R;Chow, V.T.K.; Quek, H.H.
Gene 169, 223-227, 1996

A;Title: HEP-COP, a novel human gene whose product is highly homologous to the alpha-sub
A;Reference number: JC4668; MUID:96194806; PMID:8647451
A;Accession: JC4668
A;Molecule type: mRNA
A;Residues: 1-1224 <CHO>
A;Cross-references: UNIPROT:P53621; GB:U24105; NID:g1638873; PIDN:AAB70879.1; PID:g10023
A;Experimental source: Hep3B hepatocellular carcinoma cell
R;Feurle, G.E.; Hamscher, G.; Kusiek, R.; Meyer, H.E.; Metzger, J.W.
J. Biol. Chem. 267, 22305-22309, 1992
A;Title: Identification of xenin, a xenopsin-related peptide, in the human gastric mucos
A;Reference number: A44317; MUID:93054515; PMID:1429581
A;Accession: A44317
A;Molecule type: protein
A;Residues: 1-25 <FEU>
A;Experimental source: gastric mucosa
A;Note: sequence extracted from NCBI backbone (NCBIP:117018)
A;Note: plasma levels of xenin 25, as determined by immunoassay, rise after meals and it
C;Genetics:
A;Gene: GDB:COPA; HEP-COP
A;Cross-references: GDB:4642787; OMIM:601924
A;Map position: 1q23-1q25
C;Superfamily: yeast coatomer complex alpha chain; WD repeat homology
C;Keywords: duplication; hormone; plasma; stomach
F;1-25/Product: xenin 25 #status experimental <XNP>
F;5-38/Domain: WD repeat homology <WD1>
F;17-25/Product: xenopsin-related peptide #status predicted <XRP>
F;47-80/Domain: WD repeat homology <WD2>
F;89-122/Domain: WD repeat homology <WD3>
F;131-164/Domain: WD repeat homology <WD4>
F;201-234/Domain: WD repeat homology <WD5>
F;245-278/Domain: WD repeat homology <WD6>

Query Match   7.1%;   Score 90;   DB 1;   Length 1224;
Best Local Similarity   25.9%;   Pred. No. 23;
Matches   41;   Conservative   26;   Mismatches   59;   Indels   32;   Gaps   7;

```
Qy   90 EGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIB-LKGIDFKEDGNILGHKLEYNYNSH 148
            ||: |||    |||| |||||||||||||||||||
Db  632 KGYPBVALHFVKDB---KTRFSLALECGNIBIALEAAKALDDKNCWEKLGEVALLQGNHQ 688

Qy  149 NVYIMADKQKNGIKVNF-----------KIRHNIEDGGVQ--LADHYQQNTPIGDGPVLLPD 197
            ||:  || | :||           | || ||||
Db  689 IVEMCYQRTKNFDKVSFLYLITGNLEKLRKMMKIABIRKDMSGHYQNALYLGD------- 741

Qy  198 NHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDE 235
            :|||
Db  742 --VSBRVRILKNCGQKS-----LAYLTRA--THGLDE 769
```

RESULT 6
H64102
leucine-tRNA ligase (EC 6.1.1.4) - Haemophilus influenzae (strain Rd KW20)
N;Alternate names: leucyl-tRNA synthetase
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C;Accession: H64102
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: H64102
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-861 <TIGR>
A;Cross-references: UNIPROT:P43827; GB:L42023; NID:g1573942; PIDN:AAC22581.1;
C;Genetics:
A;Gene: leuS
C;Superfamily: leucine-tRNA ligase
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match   7.0%;   Score 89.5;   DB 2;   Length 861;
Best Local Similarity   24.3%;   Pred. No. 16;
Matches   45;   Conservative   26;   Mismatches   67;   Indels   47;   Gaps   9;

Qy   49 TTGKLPVPWPTLVTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKT 108
           ||||: |||| ::|::|::||| : |||
Db   314 TGDKLPI-WVANFVLMHYGTGAVMAVPAHDQRD--FEFAQKYSLPIKQVIAPLADEEIDL 370

Qy   109 RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADK-QKNGI---KVN 164
           ||||||||||||||||||||||||||||||||||||||||||||||| ||::   :||
Db   371 TKQAFVEHGKLVNSDEFDGKNF--DGAFNG-----------------IADKLEKLGVGKRQYN 414

Qy   165 FKIRH-------------NIEDGGVQLADHYQQNTPIGDGPVLLPDNHYL-STQSAL 207
           ::|:   :|:| :|| ::|||||:||||| ::
Db   415 YRLRDWGVSRQRYWGAPIPMLTLENGDVVPA-------PMEDLPIILPBDVVMDGVKSPI 467

Qy   208 SKDPN 212
           :|||:
Db   468 NADPN 472

RESULT 7
E70390
iron-sulfur cofactor synthesis protein nifs - Aquifex aeolicus
N;Contains: L-cysteine sulfurtransferase (EC 2.8.1.-)
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: E70390
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:9816666; PMID:9537320
A;Accession: E70390
A;Molecule type: DNA
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Experimental source: strain VF5
A;Residues: 1-370 <AQF>
A;Cross-references: UNIPROT:O67152; GB:AE000720; NID:g2983529; PIDN:AAC07111.1; PID:g298
C;Genetics:
A;Gene: nifs1
C;Superfamily: nitrogen fixation protein nifs
C;Keywords: phosphoprotein; pyridoxal phosphate; sulfurtransferase
F;195/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
F;318/Active site: Cys (cysteine persulfide intermediate) #status predicted

Query Match   6.9%;   Score 88.5;   DB 2;   Length 370;
Best Local Similarity   25.4%;   Pred. No. 6.4;
Matches   49;   Conservative   30;   Mismatches   91;   Indels   23;   Gaps   8;

Qy   3 KGEELFTGVV---PILVELD---GDVNGHKF-SVSGEG----EGDATYGKLITLKFICT 49
          ||| ::|::   ::|:| || :|:| ||::|||:|:|:::|
Db   164 KGVPLLTDAVQAIGKIPIELKNISYATFSGHKFHAIKGSGFLYISDEANYEPLIVGGGQE 223

Qy   50 TGKLP-----VPWPTLVTTLSYGVQCFSRYPDHMKR-HDFFKSAMPEGYVQERTIFFKDD 103
          ||||| :|||||||||||||||||||||||| ||||||||||||||||||||||
Db   224 NGKRSGTENVVGILSLAKALEIIVSNFSRYOEQLKKLRDLFENLLLEA-LPDAQIVGKDA 282

Qy   104 GNYKTRAEV---KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNG 160
          |||:||:| |||||||||||||:| :|:|
Db   283 ERSPSISSVIMPKFFGAEIVNKLSEKGIYCSTGSACLSGEYEPNKHMLRMGFSQEKALRM 342

Qy   161 IKVNFKIRHNIED 173
          :|:||:
Db   343 VRFSFGLLNKEEE 355

RESULT 8
E84941
imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19) [imported] - Buchnera sp. (strain
C;Species: Buchnera sp.
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: E84941

R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A;Reference number: A84930; MUID:20445173; PMID:10993077
A;Accession: E84941
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-353 <STO>
A;Cross-references: GB:AP000398; GSPDB:GN00144
A;Experimental source: strain APS
C;Genetics:
A;Gene: hisB; BU102
C;Superfamily: hisB bifunctional enzyme; imidazoleglycerol-phosphate dehydratase homolog\
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match   6.9%;   Score 88;   DB 2;   Length 353;
Best Local Similarity   24.3%;   Pred. No. 6.7;
Matches   78;   Indels   42;   Gaps   10;
Conservative   33;   Mismatches

Qy   63 TLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIF---FKDDGNYKTRAEVKF----- 114
          |:|: |::|: ||: : |:|
Db   61 TESFPLQDFS--TAHLFMLSVFRS---EGVIFDDILICPHFLDDDCVCRKPKIMIBPWL 115

Qy   115 ------------EGDT---LVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKN 159
                      |||| || |:|: :|| :||| |:|:||
Db   116 DKIDLKKSYVIGDRDTDMQLSNNLKIGIKYKEDICNWLHITKY-IIKHNRYAEIIRRTK 174

Qy   160 GIKVNFKIRHNIE------DGGVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEK 214
          | |::| | |||| ||| |::|||| :|||||
Db   175 ETKVSIKVWLDLEETSKIDTGVKFFDHMLEQLSVHSGICM-----NISVQGDLDIDDHHT 229

Qy   215 -RDHMVLLG--FVTAAGITHGM 233
           :| ::|:
Db   230 IEDTGIVLGEALLQALGKGNGL 251

RESULT 9
JC4078
protective surface antigen D-15 precursor - Haemophilus influenzae (type b)
C;Species: Haemophilus influenzae
A;Variety: type b
C;Date: 30-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004
C;Accession: JC4078
R;Flack, F.S.; Loosmore, S.; Chong, P.; Thomas, W.R.
Gene 156, 97-99, 1995
A;Title: The sequencing of the 80-kDa D15 protective surface antigen of Haemophilus influ\
A;Reference number: JC4078; MUID:95255676; PMID:7737523
A;Accession: JC4078
A;Molecule type: DNA
A;Residues: 1-797 <FLA>
A;Cross-references: UNIPROT:P46024; GB:U13961; NID:g537447; PIDN:AAA85645.1; PID:g537448
A;Experimental source: type b
C;Superfamily: protective surface antigen D-15
C;Keywords: surface antigen
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-797/Product: protective surface antigen D-15 #status predicted <MAT>

Query Match   6.9%;   Score 87.5;   DB 2;   Length 797;
Best Local Similarity   22.4%;   Pred. No. 21;
Matches   49;   Conservative   34;   Mismatches   73;   Indels   63;   Gaps   12;

Qy   64 LSYGVQCFSRYPDHMKRHDF------------------FKSAMPEGVVQE---------RTI 98
          ||||||||||||||||||| ||:|||||::| |:
Db   427 IGYGTESGISYOASVKQDNFLGTGAAVSIAGTKNDYGTSVNLGYTEPYFTKDGVSLGGNV 486

Qy   99 FFKDDGNYKTRABVKFPEGDTLVNRIELKGIDFKEDGNI------LGH--------KLEYNYNS 147
          |||||||: |::|:|| |||||||||||||||||| ||| |||||||
Db   487 FFENYDNSKSDTSSNYKRTYGSNVTL-GFPVNENNSYYVGLGHTYNKIISNFALEYN--- 542

Qy   148 HNVYIMADKQK-NGIKVN------FKIRHN------IEDGGVQLADHYQQNTPIGDGPVLL 195
          |:|:::   :|| | || :||| ||| |:|| ||||
Db   543 RNLYIQSMKFKGNGIKTNDFPSFGWNYNSLNRGYFPFTKGVKAS------LG-GRVTI 593

Qy   196 P--DNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHG 232
               ||| |||::::| ||||:| ||||
Db   594 PGSDNKYYKLSADVQGFYPLDRDHLWVVSAKASAGYANG 632

RESULT 10
F64102
protective surface antigen D-15 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 18-Sep-1998
C;Accession: F64102
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
.; Bult, C.J.; Tomb, J.; Dougherty, B.A.; Merrick, J.M.; McKenney, K.; Sutton, G.; FitzHugh, W.;
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
.F.; Phillips, C.A.; Spriggs, T.; Hedblom, E.; Cotton, M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen,
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.;
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: F64102
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-808 <TIGR>
A;Cross-references: GB:L42023; TIGR:HI0917
C;Superfamily: protective surface antigen D-15
C;Keywords: surface antigen

Query Match   6.9%; Score 87.5; DB 2; Length 808;
Best Local Similarity   22.4%; Pred. No. 22;
Matches   49; Conservative   34; Mismatches   73; Indels   63; Gaps   12;

Qy   64 LSYGVQCFSRYPDHMKRHDF------------FKSAMPEGYVQE---------RTI 98
               :|:: ||||| ::::|
Db   440 IGVGTESGISYQASVKQDNFLGTGAAVSIAGTKNDYGTSVNLGVTEPYFTKDGVSLGGNV 499

Qy   99 FPKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNI---LGH---------KLEYNYNS 147
               |||||| |||:: |:|:| |:: |||:
Db   500 FFENYDNSKSDTSSNYKRTTYGSNVTL-GFPVNENNSYYVLGHTYNKISNFALEYN---- 555

Qy   148 HNVYIMADKQK-NGIKVN------FKIRHN------IEDGGVQLADHYQQNTPIGDGPVLL 195
               || ||:| |||:: ||:| : | :|
Db   556 RNLYIQSMKFKGNGIKTNDFDFDSFGWNYNSLNRGYFPTKGVKAS---------LG-GRVTI 606

Qy   196 P--DNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHG 232
               ||| |||::::| ||||:| ||||
Db   607 PGSDNKYYKLSADVQGFYPLDRDHLWVVSAKASAGYANG 645

RESULT 11
A36028
DNA-directed DNA polymerase (EC 2.7.7.7) II catalytic chain - yeast (Saccharomyces cerev
N;Alternate names: DNA-directed DNA polymerase II chain A; protein N0825; protein YNL262
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A36028; B36028; S60919; S63235; S65121
R;Morrison, A.; Araki, H.; Clark, A.B.; Hamatake, R.K.; Sugino, A.
Cell 62, 1143-1151, 1990
A;Title: A third essential DNA polymerase in Saccharomyces cerevisiae.
A;Reference number: A36028; MUID:90381771; PMID:2169349
A;Accession: A36028
A;Molecule type: DNA
A;Residues: 1-2222 <MOR>
A;Cross-references: UNIPROT:P21951; X',1218-1221 <MO2>
A;Residues: 1214-1216,'X',1218-1221 <MO2>
R;Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.
submitted to the EMBL Data Library, October 1995
A;Accession: B36028
A;Molecule type: protein
A;Residues: 1-2222 <MOR>
A;Cross-references: UNIPROT:P21951;
A;Accession: S60909
A;Reference number: S60919
A;Accession: S60919
A;Molecule type: DNA
A;Residues: 1-2221 <SEN>
A;Cross-references: EMBL:X92494; NID:g1045236; PIDN:CAA63235.1.1; PID:g1045247

Qy   196 P--DNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHG 232
               ||| |||::::| ||||:| ||||
Db   607 PGSDNKYYKLSADVQGFYPLDRDHLWVVSAKASAGYANG 645

R;Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63235
A;Accession: S63235
A;Molecule type: DNA
A;Residues: 1-2222 <SEN>
A;Cross-references: EMBL:Z71538; NID:g1302316; PIDN:CAA96169.1; PID:g1302317; GSPDB:GN00
A;Experimental source: strain S288C
R;Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.
Yeast 12, 505-514, 1996
A;Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from Sa
A;Reference number: S65111; MUID:96310631; PMID:8740425
A;Accession: S65121
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2221 <SEF>
A;Cross-references: EMBL:X92494; NID:g1045236; PIDN:CAA63235.1; PID:g1045247
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Gene: SGD:POL2; DUN2; MIPS:YNL262w
A;Cross-references: SGD:S0005206; MIPS:YNL262w
A;Map position: 14L
C;Superfamily: DNA-directed DNA polymerase II
C;Keywords: DNA binding; nucleotidyltransferase; nucleus; zinc finger

Query Match   6.9%; Score 87.5; DB 1; Length 2222;
Best Local Similarity   28.2%; Pred. No. 81;
Matches   37; Conservative   14; Mismatches   49; Indels   31; Gaps   7;

Qy   53 LPVPWP-TLVTTLSYGVQCFSRYPDHM--------KRHDFFKSAMPEGYV---QERTI 98
               ||| ||||| ||||||||
Db   883 LPKSFPETYFPTLENGKKLYLSYPCSMLNYRVHQKFTNHQYQELKDPLNYIYETHSENTI 942

Qy   99 FPKDDGNYKTR--AEVKPEGDTLVNR----------IELKGIDFKEDGNILGHKLEYNYN 146
               ||||||| ||:::
Db   943 FFEVDGPYKAMILPSSKEBGKGIKKGRYAVFNEDGSLABLKGFELKRRGEL---QLIKNFQ 999

Qy   147 S--HNVYIMAD 155
               :   ::
Db   1000 SDIFKVFLEGD 1010

RESULT 12
G81355
tRNA (uracil-5-)-methyltransferase (EC 2.1.1.35) Cj0831c [imported] - Campylobacter jeju
tRNA (uracil-5-)-methyltransferase (EC 2.1.1.35) Cj0831c [imported] - Campylobacter jeju
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: G81355
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: G81355
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-357 <PAR>
A;Cross-references: UNIPROT:Q9PP92; GB:AL139076; GB:AL111168; NID:g6968128; NID:g6968128
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: trmA; Cj0831c
C;Keywords: methyltransferase; S-adenosylmethionine

Query Match   6.8%; Score 87; DB 2; Length 357;
Best Local Similarity   24.8%; Pred. No. 8.2;
Matches   30; Conservative   18; Mismatches   39; Indels   34; Gaps   5;

Qy   79 KRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKP--EGDTLV-------NRIBLKG 127
               ||::| ||||:||
Db   14 EKHSFIKKYFKBFYTKDFKLJFASKDKHYRTRAELSFYHENDTLFYAMFDPKSKKKYIIEY 73

Qy   128 IDFKED------------GNILGHKLEYNYNSHNVYIMADKQKNGIKVNPKIRHNIE 172
               :|| :| ||| | :| || || |||

```
Db       74 LDFADEKICAFMPRLLEYLRQDNKLKEKL------FGVEFLTTKQE--LSITLLYHKNIE 125

Qy      173 D 173
Db      126 D 126

RESULT 13
JH0414
synaptogamin o-p65-B - electric ray (Discopyge ommata)
N;Alternate names: synaptic vesicle protein o-p65-B
C;Species: Discopyge ommata
C;Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 09-Jul-2004
C;Accession: JH0414; PS0223
R;Wendland, B.; Miller, K.G.; Schilling, J.; Scheller, R.H.
Neuron 6, 993-1007, 1991
A;Title: Differential expression of the p65 gene family.
A;Reference number: JH0413; MUID:91273991; PMID:2054189
A;Accession: JH0414
A;Molecule type: mRNA
A;Cross-references: UNIPROT:P24506; GB:M64276; NID:g213110; PIDN:AAA49228.1; PID:g213111
A;Experimental source: electric organ
A;Accession: PS0223
A;Molecule type: protein
A;Residues:'MLV',26-34;'XX',194-199,'X',201-206;'X',322-332,'D',334-337 <WEN1>
C;Superfamily: synaptotagmin; protein kinase C C2 region homology <KC2A>
C;Keywords: glycoprotein; membrane protein; synaptic vesicle
C;Date: hydrophobic <HYD>
F;75-101/Domain: hydrophobic <HYD>
F;153-266/Domain: protein kinase C C2 region homology <KC2A>
F;284-399/Domain: protein kinase C C2 region homology <KC2B>
F;6,46/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match             6.8%;  Score 87;  DB 2;  Length 439;
Best Local Similarity  20.1%;  Pred. No. 11;
Matches  58;  Conservative  44;  Mismatches  92;  Indels  94;  Gaps  13;

Qy       15 LVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPV-PWPTLVTTLSYGV----- 68
            ::  || : ::| |::|
Db       38 MMPIDTGDNSTRAGVPGEGKND-VFEKLKEKFMNELQKIPILPPWALIAIAIVSGLLLLTC 96

Qy       69 -----QCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG------- 116
                 | |::| |   ||  :|:::
Db       97 CLCICKKCCCCKKKNKEKGKGK---------KNDINMK---DVKGSSGNQDDD 138

Qy      117 --DTLVNRIELKGIDFKEDGNI--LGHKLEYNYNSH----------------- 148
               :|:: |  ::::|
Db      139 DAETGLTEGEDKEEEAKEEEKLGKIQFSLDYDFQANQLTVGIIQAAELPALDMGGTSDPY 198

Qy      149 -NVYIMADKQKN-GIKVN--------------FKIRHNIEDGGVQLA-------DHYQQNTP 187
             :||:|||:||  |:::               ||:|  | :| ||
Db      199 VKVFLLPDKKKKYETKVQKKTLNPTFNESFVFKVPYQ-ELGGKTLMMAVYDFDRFSKHDC 257

Qy      188 IGDGPVLLPD------NHYLSTQSALSKDPNEKRDHMVLLGFVTAAG 228
            | ::   | |      :|| ::|:|  |::|
Db      258 IQQVTVLMTKVDLGQQLEEWRDLESAEKEEPEKLGDICTSLRYVPTAG 305

RESULT 14
S53488
water-stress-inducible protein DS2 - Chaco potato
C;Species: Solanum chacoense (Chaco potato)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S53488
R;Silhavy, D.; Hutvagner, G.; Barta, E.; Banfalvi, Z.
Plant Mol. Biol. 27, 587-595, 1995
A;Title: Isolation and characterization of a water-stress-inducible cDNA clone from Sola
A;Accession: S53488
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-263 <SIL>
A;Cross-references: UNIPROT:Q41300; GB:U12439; NID:g607904; PIDN:AAA86052.1; PID:g607905
```

```
Query Match             6.8%;  Score 86.5;  DB 2;  Length 263;
Best Local Similarity  19.8%;  Pred. No. 6.1;
Matches  50;  Conservative  14;  Mismatches  89;  Indels  99;  Gaps  8;

Qy       31 GEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLSYGVQCFSRYPDHMKRHDFFKSAMPE 90
            |: ||::|||||
Db       31 ----------------------------------------------SYGDDTYG------ 51

Qy       91 GYVQERTIFFKDDGNYKTRAE----------VKFEGDTLVNRIEL-KGIDFK------- 131
                ||| :|   |
Db       52 ----EKTTTFGDDNKYGEKTSYGGDNTYGEKTSYGKGDDNKYGEKTSY 107

Qy      132 --EDGNILGHKLEY---NYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLADHYQQNT 186
                :| |::                |||:||:| |::||:|:|:   |
Db      108 GEGDDNKYGEKTSYGDSGYGEKPSYGGDDNKYGEKTSYGNEEGGYGGVGETTNYEENE 167

Qy      187 ------------------------------PIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVL 220
                                          |:  | |:|
Db      168 SETKTSEDYKEEKKHKHKHLEIGGLGAVRAGAFALHEKH------KAEKDPENAHKHKIE 221

Qy      221 LGFVTAARGITHG 232
            :|
Db      222 EGIAAAAIGAG 233

RESULT 15
AC0582
leucyl-tRNA synthetase [imported] - Salmonella enterica subsp. enterica serovar Typhi (st
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AC0582
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serova
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC0582
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-860 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05125.1; PID:g16501899; GSPDB:GN00176
C;Genetics:
A;Gene: STY0679
C;Superfamily: leucine-tRNA ligase

Query Match             6.8%;  Score 86.5;  DB 2;  Length 860;
Best Local Similarity  23.7%;  Pred. No. 29;
Matches  42;  Conservative  20;  Mismatches  82;  Indels  33;  Gaps  6;

Qy       49 TTGKLPVPWPTLVTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKT 108
            |||  |    |  ::|
Db      314 TGEEIPV-WAANFVLMEYGTGAVMAVPGHDQRD--YEFASKYGLTIKPVILAADGSEPDL 370

Qy      109 RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIR 168
                :   |::|:    |
Db      371 SEQALTEKGVLFNSGEFDGLAFEAAFNAIADKL---------AEKGVGERKVNYRLR 418

Qy      169 HNIEDGGVQ-----------LADHYQQNTPIGDGPVLLPDNHYL-STQSALSKDP 211
            :||:|
Db      419 ----DWGVSRQRYWGAPIPMVTLEDGTVLPTPEDQLPVILPBDVVMDGITSPIKADP 471

Search completed: September 16, 2005, 17:21:02
Job time : 25.5 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:   September 16, 2005, 16:55:06 ; Search time 115 Seconds
                                      (without alignments)
                               1059.781 Million cell updates/sec

Title:          US-10-757-624-3
Perfect score:  1276
Sequence:       1 MSKGEBLFTGVPILVELDG..........VLLGFVTAAGITHGMDELYK 238

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:      1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   UniProt_03:*
             1: uniprot_sprot:*
             2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1258	98.6	238	2	Q8GHE2	Q8ghe2 azotobacter
2	1254	98.3	238	1	GFP_AEQVI	P42212 aequorea vi
3	1254	98.3	238	2	Q71RY9	Q71ry9 azotobacter
4	1253	98.2	238	2	Q8GHE4	Q8ghe4 azomonas ag
5	1252	98.1	238	2	Q8GHE3	Q8ghe3 azotobacter
6	1247	97.7	238	2	Q93125	Q93125 aequorea vi
7	1212	95.0	238	2	Q17105	Q17105 aequorea vi
8	1197	93.8	238	2	Q17106	Q17106 aequorea vi
9	1184	92.8	238	2	Q6YGZ0	Q6ygz0 aequorea co
10	1105	86.6	238	2	Q8WTC6	Q8wtc6 aequorea ma
11	1101	86.3	238	2	Q8WP95	Q8wp95 aequorea ma
12	1097	86.0	238	2	Q8WTC4	Q8wtc4 aequorea ma
13	1095	85.8	238	2	Q8WTD0	Q8wtd0 aequorea ma
14	1094	85.7	238	2	Q8WTC8	Q8wtc8 aequorea ma
15	1094	85.7	238	2	Q8WTC9	Q8wtc9 aequorea ma
16	1092	85.6	238	2	Q8WTC7	Q8wtc7 aequorea ma
17	1090	85.4	238	2	Q8WTC5	Q8wtc5 aequorea ma
18	625	49.0	234	2	Q6RYS7	Q6rys7 phialidium
19	473	37.1	225	2	Q6RYS5	Q6rys5 anthomedusa
20	436	34.2	262	2	Q6RYS6	Q6rys6 anthomedusa
21	253.5	19.9	226	2	Q8T6U0	Q8t6u0 dendronepht
22	252.5	19.8	225	2	Q7Z0W5	Q7z0w5 montastraea
23	252.5	19.8	225	2	Q963F5	Q963f5 montastraea
24	250	19.6	225	2	Q6RYS4	Q6rys4 anthomedusa
25	250	19.6	232	2	Q7Z0W9	Q7z0w9 montastraea
26	248.5	19.5	225	2	Q66PW1	Q66pw1 scolymia cu
27	244.5	19.2	230	2	Q9U6Y3	Q9u6y3 clavularia
28	244.5	19.2	266	2	Q6USK3	Q6usk3 montastraea
29	242.5	19.0	225	2	Q8I6J8	Q8i6j8 trachyphyll
30	240	18.8	225	2	Q8MU48	Q8mu48 montastraea
31	238	18.7	224	2		
32	238	18.7	225	2	Q66ND3	Q66nd3 montastraea
33	237.5	18.6	225	2	Q8T5F1	Q8t5f1 montastraea
34	236	18.5	225	2	Q7Z0W4	Q7z0w4 montastraea
35	222.5	17.4	221	2	Q66PV1	Q66pv1 acropora te
36	222.5	17.4	227	2	Q66ND4	Q66nd4 montastraea
37	222	17.4	231	2	Q66PV5	Q66pv5 acropora mi
38	220.5	17.3	223	2	Q6R8F5	Q6r8f5 astrangia l
39	220	17.2	227	2	Q66ND2	Q66nd2 montastraea
40	220	17.2	227	2	Q66ND5	Q66nd5 montastraea
41	219.5	17.2	221	2	Q66PU8	Q66pu8 acropora ac
42	218.5	17.1	259	2	Q8MMA2	Q8mma2 agaricia fr
43	218	17.1	239	2	Q8MMA1	Q8mma1 agaricia ag
44	217	17.0	227	2	Q7Z0W6	Q7z0w6 montastraea
45	216.5	17.0	234	2	Q7Z0W7	Q7z0w7 montastraea

ALIGNMENTS

```
RESULT 1
Q8GHE2
ID   Q8GHE2       PRELIMINARY;       PRT;   238 AA.
AC   Q8GHE2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescence protein.
GN   Name=2289Gfp;
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM2289;
RA   Koranyi P., Berenyi M., Burg K.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF324408; AAN86140.1; -.
     HSSP; P42212; 1GFL.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
SQ   SEQUENCE   238 AA;  26914 MW;  F84840F1F9064018 CRC64;

Query Match          98.6%;  Score 1258; DB 2;   Length 238;
Best Local Similarity  98.7%;  Pred. No. 2.7e-94;
Matches 235;  Conservative  0; Mismatches  3;  Indels  0;  Gaps  0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60

Qy  61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNPKIRHNIEDGGVQLAD 180
        |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 2
GFP_AEQVI
ID   GFP_AEQVI    STANDARD;        PRT;    238 AA.
AC   P42212; Q17104; Q27903;
```

```
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Green fluorescent protein.
GN   Name=GFP;
OS   Aequorea victoria (Jellyfish).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=6100;
RN   [1]
RP   SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX   MEDLINE=9217527; PubMed=1342277; DOI=10.1016/0378-1119(92)90691-H;
RA   Prasher D.C., Eckenrode V.K., Ward W.W., Prendergast F.G.,
RA   Cormier M.J.;
RT   "Primary structure of the Aequorea victoria green-fluorescent
RT   protein.";
RL   Gene 111:229-233(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94185810; PubMed=8137953; DOI=10.1016/0014-5793(94)80472-9;
RA   Inouye S., Tsuji F.I.;
RT   "Aequorea green fluorescent protein. Expression of the gene and
RT   fluorescence characteristics of the recombinant protein.";
RL   FEBS Lett. 341:277-280(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97299832; PubMed=9154981; DOI=10.1023/A:1005740823703;
RA   Rouwendal G.J.A., Mendes O., Wolbert E.J.H., de Boer A.D.;
RT   "Enhanced expression in tobacco of the gene encoding green fluorescent
RT   protein by modification of its codon usage.";
RL   Plant Mol. Biol. 33:989-999(1997).
RN   [4]
RP   CHROMOPHORE.
RX   MEDLINE=93192221; PubMed=8448132;
RA   Cody C.W., Prasher D.C., Westler W.M., Prendergast F.G., Ward W.W.;
RT   "Chemical structure of the hexapeptide chromophore of the Aequorea
RT   green-fluorescent protein.";
RL   Biochemistry 32:1212-1218(1993).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=96355665; PubMed=8703075;
RA   Ormoe M., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y.,
RA   Remington S.J.;
RT   "Crystal structure of the Aequorea victoria green fluorescent
RT   protein.";
RL   Science 273:1392-1395(1996).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=98294543; PubMed=9631087;
RA   Yang F., Moss L.G., Phillips G.N. Jr.;
RT   "The molecular structure of green fluorescent protein.";
RL   Nat. Biotechnol. 14:1246-1251(1996).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF MUTANT WITH YELLOW EMMISSION.
RX   MEDLINE=98455509; PubMed=9782051; DOI=10.1016/S0969-2126(98)00127-0;
RA   Wachter R.M., Elsliger M.A., Kallio K., Hanson G.T., Remington S.J.;
RT   "Structural basis of spectral shifts in the yellow-emission variants
RT   of green fluorescent protein.";
RL   Structure 6:1267-1277(1998).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX   MEDLINE=99238303; PubMed=10220315; DOI=10.1021/bi9902182;
RA   Elsliger M.A., Wachter R.M., Hanson G.T., Kallio K., Remington S.J.;
RT   "Structural and spectral response of green fluorescent protein
RT   variants to changes in pH.";
RL   Biochemistry 38:5296-5301(1999).
CC   -!- FUNCTION: Energy-transfer acceptor. Its role is to transduce the
CC       blue chemiluminescence of the protein aequorin into green
CC       fluorescent light by energy transfer. Fluoresces in vivo upon
CC       receiving energy from the Ca(2+)-activated photoprotein aequorin.
CC       Absorbs light maximally at 395 nm and exhibits a smaller
CC       absorbance peak at 470 nm. The fluorescence emission spectrum
CC       peaks at 509 nm with a shoulder at 540 nm.
CC   -!- SUBUNIT: Monomer.
CC   -!- TISSUE SPECIFICITY: Photocytes.
CC   -!- PTM: Contains a covalently attached chromophore, which is composed
CC       of modified amino acid residues. The chromophore is formed upon
CC       cyclization of the residues Ser-dehydroTyr-Gly.
CC   -!- BIOTECHNOLOGY: Has become a useful and ubiquitous tool for making
CC       chimeric proteins of GFP linked to other proteins where it
CC       functions as a fluorescent protein tag. GFP tolerates N-and C-
CC       terminal fusion to a broad variety of proteins. It has been
CC       expressed in bacteria, yeast, slime mold, plants, Drosophila,
CC       zebrafish, and in mammalian cells. As a noninvasive fluorescent
CC       marker in living cells, it allows for a wide range of applications
CC       where it may function as a cell lineage tracer, reporter of gene
CC       expression, or as a measure of protein-protein interactions.
CC   -!- DATABASE: NAME=Protein Spotlight; NOTE=Issue 11 of June 2001;
CC       WWW="http://www.expasy.org/spotlight/articles/sptlc011.html".
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M62654; AAA27722.1; -.
DR   EMBL; M62653; AAA27721.1; -.
DR   EMBL; L29345; AAA58246.1; -.
DR   EMBL; X96418; CAA65278.1; -.
DR   PIR; JS0692; JQ1514.
DR   PDB; 1B9C; X-ray; A/B/C/D=1-238.
DR   PDB; 1BFP; X-ray; @=1-238.
DR   PDB; 1C4F; X-ray; A=1-238.
DR   PDB; 1CV7; X-ray; A=1-228.
DR   PDB; 1EMA; X-ray; @=-.
DR   PDB; 1EMB; X-ray; @=1-238.
DR   PDB; 1EMC; X-ray; A/B/C/D=1-238.
DR   PDB; 1EME; X-ray; @=1-238.
DR   PDB; 1EMF; X-ray; @=1-238.
DR   PDB; 1EMG; X-ray; A=1-238.
DR   PDB; 1EMK; X-ray; @=1-238.
DR   PDB; 1EML; X-ray; @=1-238.
DR   PDB; 1EMM; X-ray; @=1-238.
DR   PDB; 1F09; X-ray; A=1-238.
DR   PDB; 1F0B; X-ray; A=1-238.
DR   PDB; 1GFL; X-ray; A/B=1-238.
DR   PDB; 1H6R; X-ray; A/B/C=1-238.
DR   PDB; 1HCJ; X-ray; A/B/C/D=1-238.
DR   PDB; 1HUY; X-ray; A=1-238.
DR   PDB; 1JBY; X-ray; A=1-238.
DR   PDB; 1JBZ; X-ray; A=1-238.
DR   PDB; 1JC0; X-ray; A/B/C=1-238.
DR   PDB; 1JC1; X-ray; A/B/C=1-238.
DR   PDB; 1KP5; X-ray; A/B=1-237.
DR   PDB; 1KYP; X-ray; A=1-238.
DR   PDB; 1KYR; X-ray; A=1-238.
DR   PDB; 1KYS; X-ray; A=1-238.
DR   PDB; 1MYW; X-ray; A=1-238.
DR   PDB; 1Q4A; X-ray; A=1-238.
DR   PDB; 1Q4B; X-ray; A=1-238.
DR   PDB; 1Q4C; X-ray; A=1-238.
DR   PDB; 1Q4D; X-ray; A=1-238.
DR   PDB; 1Q4E; X-ray; A=1-238.
DR   PDB; 1Q73; X-ray; A=1-238.
DR   PDB; 1QXT; X-ray; A=1-229.
DR   PDB; 1QY3; X-ray; A=1-229.
DR   PDB; 1QYF; X-ray; A=1-229.
DR   PDB; 1QYO; X-ray; A=1-238.
DR   PDB; 1QYQ; X-ray; A=1-238.
DR   PDB; 1YFP; X-ray; A/B=3-228.
DR   PDB; 2EMD; X-ray; @=1-238.
DR   PDB; 2EMN; X-ray; @=1-238.
DR   PDB; 2EMO; X-ray; @=1-238.
```

RESULT 3
Q71RY9
ID   Q71RY9   PRELIMINARY;   PRT;   238 AA.
AC   Q71RY9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Green fluorescence protein.
GN   Name=289Gfp;
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RC   SEQUENCE FROM N.A.
RC   STRAIN=CM289;
RA   Koranyi P., Berenyi M., Burg K.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF324407; AAN86139.1; -.
DR   HSSP; P42212; 1B9C.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
SQ   SEQUENCE   238 AA; 26886 MW; EA5A6F21FBFB6E05 CRC64;

Query Match   98.3%; Score 1254; DB 2; Length 238;
Best Local Similarity   98.3%; Pred. No. 5.6e-94;
Matches 234; Conservative   1; Mismatches   3; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 4
Q8GHE4
ID   Q8GHE4   PRELIMINARY;   PRT;   238 AA.
AC   Q8GHE4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescence protein.
GN   Name=375Gfp;
OS   Azomonas agilis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azomonas.
OX   NCBI_TaxID=116849;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Koranyi P., Berenyi M., Burg K.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF324405; AAN86137.1; -.
DR   HSSP; P42212; 1BFP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.

DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
KW   3D-structure; Direct protein sequencing; Luminescence.
FT   CROSSLNK   65   67   5-imidazolinone (Ser-Gly).
FT   MOD_RES   66   66   (Z)-2,3-didehydrotyrosine.
FT   VARIANT   100   100   T -> Y.
FT   VARIANT   108   108   F -> S.
FT   VARIANT   141   141   L -> M.
FT   VARIANT   219   219   V -> I.
FT   CONFLICT   25   25   S -> G (in Ref. 3).
FT   CONFLICT   80   80   H -> Q (in Ref. 2).
FT   CONFLICT   157   157   Q -> P (in Ref. 3).
FT   CONFLICT   172   172   Q -> Q (in Ref. 2).
FT   CONFLICT   172   172   E -> K (in Ref. 2).
FT   TURN   2   2
FT   HELIX   4   8
FT   STRAND   11   22
FT   TURN   23   24
FT   STRAND   25   36
FT   HELIX   37   39
FT   TURN   40   40
FT   STRAND   41   48
FT   TURN   49   50
FT   HELIX   57   60
FT   TURN   61   63
FT   HELIX   69   71
FT   STRAND   73   73
FT   HELIX   76   81
FT   STRAND   83   86
FT   HELIX   87   90
FT   STRAND   92   100
FT   TURN   101   102
FT   STRAND   105   115
FT   TURN   116   117
FT   STRAND   118   128
FT   TURN   132   133
FT   STRAND   135   139
FT   TURN   141   141
FT   STRAND   148   155
FT   HELIX   156   158
FT   TURN   159   159
FT   STRAND   160   171
FT   TURN   172   173
FT   STRAND   176   187
FT   TURN   199   208
FT   TURN   211   212
FT   STRAND   217   227
SQ   SEQUENCE   238 AA; 26886 MW; EA5A6F21FBFB6E05 CRC64;

Query Match   98.3%; Score 1254; DB 1; Length 238;
Best Local Similarity   98.3%; Pred. No. 5.6e-94;
Matches 234; Conservative   1; Mismatches   3; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              |||||||||||||||||||||||||||||||||||||||||| :|||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
              ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
SQ   SEQUENCE   238 AA;   26902 MW;   15FEB9C5B4F6B89 CRC64;

Query Match
Best Local Similarity   98.2%;   98.3%;   Score 1253; DB 2; Length 238;
Matches 234;   Conservative   0;   Mismatches   4;   Indels   0;   Gaps   0;   Pred. No. 6.8e-94;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||.|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        |||||||.||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NRIELKGTDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 5
Q8GHE3
ID   Q8GHE3   PRELIMINARY;   PRT;   238 AA.
AC   Q8GHE3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescence protein.
GN   Name=85G&p;
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Koranyi P., Berenyi M., Burg K.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF324406; AAN86138.1; -.
DR   HSSP; P42212; 1GFL.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
SQ   SEQUENCE   238 AA;   26887 MW;   E0E1616BD2AF6188 CRC64;

Query Match
Best Local Similarity   98.1%;   98.3%;   Score 1252; DB 2; Length 238;
Matches 234;   Conservative   0;   Mismatches   4;   Indels   0;   Gaps   0;   Pred. No. 8.2e-94;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        |||||||||||||||||||||||||||||||||||||||||||.||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLTFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||.|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 6
Q93125
ID   Q93125   PRELIMINARY;   PRT;   238 AA.
AC   Q93125;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescent protein mutant 3.
GN   Name=GFP;
OS   Aequorea victoria (Jellyfish).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=6100;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96305137; PubMed=8707053; DOI=10.1016/0378-1119(95)00685-0;
RA   Cormack B.P., Valdivia R.H., Falkow S.;
RT   "FACS-optimized mutants of the green fluorescent protein (GFP).";
RL   Gene 173:33-38(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97195776; PubMed=9043107;
RA   Cormack B.P., Bertram G., Egerton M., Gow N.A.R., Falkow S.,
RA   Brown A.J.P.;
RT   "Yeast-enhanced green fluorescent protein (yEGFP)a reporter of gene
RT   expression in Candida albicans.";
RL   Microbiology 143:303-311(1997).
DR   EMBL; U73901; AAB18957.1; -.
DR   PDB; 2YFP; X-ray; A=1-238.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
SQ   SEQUENCE   238 AA;   26840 MW;   A2862809A9DEA60 CRC64;

Query Match
Best Local Similarity   97.7%;   97.5%;   Score 1247; DB 2; Length 238;
Matches 232;   Conservative   2;   Mismatches   4;   Indels   0;   Gaps   0;   Pred. No. 2.1e-93;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||..|||||||.|||||.||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFGYGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||.||||||||||||||
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 7
Q17105
ID   Q17105   PRELIMINARY;   PRT;   238 AA.
AC   Q17105;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescent protein (fragment).
GN   Name=GFP;
OS   Aequorea victoria (Jellyfish).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=6100;

RN [1]
RP SEQUENCE FROM N.A.
RA Watkins J.N., Campbell A.K.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR EMBL; X83959; CAA58789.1; -.
DR PIR; JS0692; JQ1514.
DR HSSP; P42212; 1GFL.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR011584; GFP_related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
FT NON_TER   238   238
SQ SEQUENCE   238 AA; 26950 MW; 26B2B450B748B44 CRC64;

Query Match   95.0%; Score 1212; DB 2; Length 238;
Best Local Similarity   94.1%; Pred. No. 1.5e-90;
Matches 224; Conservative   6; Mismatches   8; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           |||||||||||| |||||||||||| |||||||||||||||| |||||||||||||||||
Db   1 MSKGEELFTGVVPVLVELDGDVNGQKFSVSGEGEGDATYGKLTLNFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           |||:||||||||||||||| ||| |||||||||||||||:||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFYKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           ||||||||||||||||||||:|||||||||||| ||:||||||||||||||||:||:||||
Db   121 NRIELKGIDFKEDGNILGHKMEYNYNSHNVYIMGDKPKNGIKVNFKIRHNIKDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
           |||||||||||||||||||||||||||||||||||||||||:||||| |||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMILLEFVTAARITHGMDELYK 238

RESULT 8
Q17106
ID Q17106   PRELIMINARY;   PRT;   238 AA.
AC Q17106;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein (Fragment).
GN Name=GFP;
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
RN [1]
RP SEQUENCE FROM N.A.
RA Watkins J.N., Campbell A.K.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR EMBL; X83960; CAA58790.1; -.
DR PIR; JS0692; JQ1514.
DR HSSP; P42212; 1BFP.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR011584; GFP_related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
FT NON_TER   238   238
SQ SEQUENCE   238 AA; 26867 MW; BD4648262D8EABD4 CRC64;

Query Match   93.8%; Score 1197; DB 2; Length 238;
Best Local Similarity   93.3%; Pred. No. 2.4e-89;
Matches 222; Conservative   7; Mismatches   9; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGQKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db   61 VTTFSYGVQCFSRYPDHMKQHDFLKSAMPEGYVQERTIFYKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
Db   121 NRIELKGIDFKEDGNILGHKMEYNYNSHNVYIMGDKPKNGIKVNFKIRHNIKDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSQDPHGKRDHMVLLEFVTSAGITHGMDELYK 238

RESULT 9
Q6YGZ0
ID Q6YGZ0   PRELIMINARY;   PRT;   238 AA.
AC Q6YGZ0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Green fluorescent protein.
OS Aequorea coerulescens (belt jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=210840;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22726112; PubMed=12693991; DOI=10.1042/BJ20021966;
RA Gurskaya N.G., Fradkov A.F., Pounkova N.I., Staroverov D.B.,
RA Bulina M.B., Yanushevich Y.G., Labas Y.A., Lukyanov S., Lukyanov K.A.;
RT "A colourless green fluorescent protein homologue from the non-
RT fluorescent hydromedusa Aequorea coerulescens and its fluorescent
RT mutants.";
RL Biochem. J. 373:403-408(2003).
DR EMBL; AY151052; AAN41637.1; -.
DR HSSP; P42212; 1B9C.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR011584; GFP_related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
DR ProDom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE   238 AA; 26896 MW; DE72EDB7BD9F9FE CRC64;

Query Match   92.8%; Score 1184; DB 2; Length 238;
Best Local Similarity   91.2%; Pred. No. 2.8e-88;
Matches 217; Conservative   9; Mismatches   12; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           ||||.||||||||||.||.||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGAELFTGVVPILIELNGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db   61 VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYIQERTIFFKDDGNYKSRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
Db   121 NRIELTGTDFKEDGNILGNKMEYNYNAHNVYIMTDKAKNGIKVNFKIRHNIEDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSTLSKDPNEKRDHMIYFEFVTAAAITHGMDELYK 238

RESULT 10
Q8WTC6
ID Q8WTC6   PRELIMINARY;   PRT;   238 AA.
AC Q8WTC6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

```
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescent protein.
GN   Name=GFP;
OS   Aequorea macrodactyla.
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=147615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA   Li S.J., Xia N.S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF435431; AAL33916.1; -.
DR   HSSP; P42212; 1KYP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
SQ   SEQUENCE   238 AA;  27015 MW;  6B8FD75E88926903 CRC64;

Query Match             86.6%;  Score 1105;  DB 2;  Length 238;
Best Local Similarity   83.6%;  Pred. No. 7.4e-82;
Matches  199;  Conservative  19;  Mismatches  20;  Indels  0;  Gaps  0;

Qy     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
         |||||||||| || |:||||||:|||||| |||||||:|||||:||||||||||||||||
Db     1 MSKGEELFTGIVPVLIELDGDVHGHKFSVRGEGEGDADYGKLEIKFICTTGKLPVPWPTL 60

Qy    61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||:||| ||| ||::||: ||||||||||||:|||||| |||:||||| |||||||||
Db    61 VTTFSYGIQCFARYPEHMKMNDFFKSAMPEGYIQERTIFQDDGKYKTRGEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         |||||||:||||||||||||||||:||||||||| || :|:||||||||||||: ||||||
Db   121 NRIELKGMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
         |||.|.|:||||||:| |||||||:|:|||.||.|||||| | : ..| ||||||||||
Db   181 HYQTNVPLGDGPVLIPINHYLSTQTAISKDRNETRDHMVLEFFSACGHTHGMDELYK 238

RESULT 11
Q8WP95
ID   Q8WP95      PRELIMINARY;        PRT;     238 AA.
AC   Q8WP95;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Green fluorescent protein.
GN   Name=GFPxm;
OS   Aequorea macrodactyla.
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=147615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Qin Y.X., Pang S.Q.,
RA   Li S.J., Xia N.S.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY013824; AAK02062.1; -.
DR   EMBL; AY013821; AAK02059.1; -.
DR   HSSP; P42212; 1BFP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
SQ   SEQUENCE   238 AA;  27049 MW;  8185D0E5E529012B CRC64;

Query Match             86.3%;  Score 1101;  DB 2;  Length 238;
Best Local Similarity   83.2%;  Pred. No. 1.6e-81;
Matches  198;  Conservative  19;  Mismatches  21;  Indels  0;  Gaps  0;

Qy     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
         |||||||||| || |:||||||:|||||| |||||||:|||||:||||||||||||||||
Db     1 MSKGEELFTGIVPVLIELDGDVHGHKFSVRGEGEGDADYGKLEIKFICTTGKLPVPWPTL 60

Qy    61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||:||| ||| ||::||: ||||||||||||:|||||| |||:||||| |||||||||
Db    61 VTTFSYGIQCFARYPEHMKMNDFFKSAMPEGYIQERTIFQDDGKYKTRGEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         |||||||:||||||||||||||||:||||||||| || :|:||||||||||||: ||||||
Db   121 NRIELKGMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
         |||.|.|:||||||:| |||||||:|:|||.||.|||||| | : ..| ||||||||||
Db   181 HYQTNVPLGDGPVLIPINHYLSTQTAISKDRNETRDHMVLEFFSACGHTHGMDELYK 238

RESULT 12
Q8WTC4
ID   Q8WTC4      PRELIMINARY;        PRT;     238 AA.
AC   Q8WTC4;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescent protein.
GN   Name=GFP;
OS   Aequorea macrodactyla.
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=147615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA   Li S.J., Xia N.S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF435433; AAL33918.1; -.
DR   HSSP; P42212; 1KYP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
SQ   SEQUENCE   238 AA;  26956 MW;  75521BAF0CEBA73A CRC64;

Query Match             86.0%;  Score 1097;  DB 2;  Length 238;
Best Local Similarity   84.0%;  Pred. No. 3.3e-81;
Matches  200;  Conservative  16;  Mismatches  22;  Indels  0;  Gaps  0;

Qy     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
         ||||||||||||||||||||||:||||||||||||||:|||||:||||||||||||||||
Db     1 MSKGEELFTGVVPILVELDGDVHGHKFSVSGEGEGDADYGKLEIKFICTTGKLPVPWPTL 60

Qy    61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||| || :|| ||::||: ||||||||||||:|||||| |||:||||| |||||||||
Db    61 VTTLGYGILCFARYPEHMKMNDFFKSAMPEGYIQERTIFQDDGKYKTRGEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         |||||||:||||||||||||||||:||||||||| || :|:||||||||||||: ||||||
Db   121 NRIELKGMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
         |||.|.|:||||||:| |||||||:|:|||.||.|||||| | : ..| ||||||||||
Db   181 HYQTNVPLGDGPVLIPINHYLSTQTAISKDRNETRDHMVLEFFSACGHTHGMDELYK 238

RESULT 13
Q8WTD0
ID   Q8WTD0      PRELIMINARY;        PRT;     238 AA.
AC   Q8WTD0;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescent protein.
GN   Name=GFP;
OS   Aequorea macrodactyla.
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=147615;
```

RN   [1]
RP   SEQUENCE FROM N.A.
RA   Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA   Li S.J., Xia N.S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF435427; AAL33912.1; -.
DR   HSSP; P42212; 1KYP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
SQ   SEQUENCE   238 AA;   26997 MW;   5F80A192173CB84D CRC64;

Query Match   85.8%; Score 1095; DB 2; Length 238;
Best Local Similarity   82.8%; Pred. No. 4.8e-81;
Matches 197; Conservative 19; Mismatches 22; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           |||||||||||:||:|||||||||||||||||||||||||||:|||||||||||||||||
Db   1 MSKGEELFTGIVPVLIELDGDVHGHKFSVRGEGEGDADYGKLEIKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRABVKFEGDTLV 120
           ||:||||||||:|:||||||||| :|:||::||||||||||:|||||||||||||||||
Db   61 VTTLGYGIQCFARYPEHMKMNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           |||||||||||||||||||||||||:|:||||||||||||||||||||||||||||||||
Db   121 NRIELKGMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
           ||| :|||||||||::::|||||| :: :|| :|||:||||:||:|:||||||||||
Db   181 HYQTNVPLGDGPVLIPINHYLSLQTAISKDRNETRDHMVFLEFFSACGHTHGMDELYK 238

RESULT 15
Q8WTC9
ID   Q8WTC9   PRELIMINARY;   PRT;   238 AA.
AC   Q8WTC9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescent protein.
GN   Name=GFP;
OS   Aequorea macrodactyla.
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=147615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA   Li S.J., Xia N.S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF435428; AAL33913.1; -.
DR   HSSP; P42212; 1KYP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
SQ   SEQUENCE   238 AA;   27031 MW;   5F80A19C19DC584D CRC64;

Query Match   85.7%; Score 1094; DB 2; Length 238;
Best Local Similarity   82.8%; Pred. No. 5.8e-81;
Matches 197; Conservative 19; Mismatches 22; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           |||||||||||:||:|||||||||||||||||||||||||||:|||||||||||||||||
Db   1 MSKGEELFTGIVPVLIELDGDVHGHKFSVRGEGEGDADYGKLEIKFICTTGKLPVPWPTL 60

Qy   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRABVKFEGDTLV 120
           ||:||||||||:|:||||||||| :|:||::||||||||||:|||||||||||||||||
Db   61 VTTLGYGIQCFARYPEHMKMNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEGGGVQLAD 180
           |||||||||||||||||||||||||:|:||||||||||||||||||||||||||||||||
Db   121 NRIELKGMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
           ||| :|||||||||::::|||||| :: :|| :|||:||||:||:|:||||||||||
Db   181 HYQTNVPLGDGPVLIPINHYLSFQTAISKDRNETRDHMVFLEFFSACGHTHGMDELYK 238

Search completed: September 16, 2005, 17:20:07
Job time : 117 secs

THIS PAGE BLANK (USPTO)

```
          GenCore version 5.1.6
   Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      September 16, 2005, 16:54:01 ; Search time 117 Seconds
                            (without alignments)
                            786.743 Million cell updates/sec

Title:         US-10-757-624-4
Perfect score: 1276
Sequence:      1 MSKGEELFTGVVPILVELDG........VLLEFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:      2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:    2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    A_Geneseq_16Dec04:*
              1: geneseqp1980s:*
              2: geneseqp1990s:*
              3: geneseqp2000s:*
              4: geneseqp2001s:*
              5: geneseqp2002s:*
              6: geneseqp2003as:*
              7: geneseqp2003bs:*
              8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                   SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1276	100.0	238	6	ADA25196	Ada25196 Green flu
2	1272	99.7	238	6	ADA25218	Ada25218 Green flu
3	1270	99.5	238	2	AAW76106	Aaw76106 A. victor
4	1268	99.4	238	6	ADA25214	Ada25214 Green flu
5	1267	99.3	238	6	ABR44424	Abr44424 F64L-Y66H
6	1266	99.2	238	2	AAW22101	Aaw22101 Aequorea
7	1266	99.2	238	2	AAW65078	Aaw65078 A. victor
8	1266	99.2	238	2	AAW96328	Aaw96328 Humanised
9	1266	99.2	238	5	AAE16042	Aae16042 Aequorea
10	1266	99.2	238	6	ABG76008	Abg76008 Jellyfish
11	1266	99.2	238	6	ADA25212	Ada25212 Green flu
12	1266	99.2	238	6	ADA25224	Ada25224 Green flu
13	1266	99.2	247	2	AAW96329	Aaw96329 Humanised
14	1266	99.2	501	2	AAW31879	Aaw31879 GFP varia
15	1265	99.1	238	6	ADA25195	Ada25195 Green flu
16	1264	99.1	238	6	ABG32368	Abg32368 Aequorea
17	1264	99.1	238	6	AAE34996	Aae34996 Aequorea
18	1264	99.1	238	6	ADA25219	Ada25219 Green flu
19	1262	98.9	238	6	AAW05304	Aaw05304 Green flu
20	1262	98.9	238	2	AAW22097	Aaw22097 Aequorea
21	1262	98.9	238	2	AAW24232	Aaw24232 Aequorea
22	1262	98.9	238	2	AAW76105	Aaw76105 A. victor
23	1262	98.9	238	2	AAW40479	Aaw40479 A. victor
24	1262	98.9	238	2	AAW76371	Aaw76371 A. victor
25	1262	98.9	238	4	AAB73552	Aab73552 Wild-type
26		98.9	1262	238	5	AAE16038 Aae16038 Aequorea
27		98.9	1262	238	6	ABG32365 Abg32365 Aequorea
28		98.9	1262	238	6	ABG75980 Abg75980 Jellyfish
29		98.9	1262	238	6	AAE34999 Aae34999 Aequorea
30		98.9	1262	238	6	AAE34985 Aae34985 Aequorea
31		98.9	1262	238	6	ABR44423 Abr44423 Wild-type
32		98.9	1262	238	6	ADA25217 Ada25217 Green flu
33		98.9	1262	238	6	ADA25194 Ada25194 Aequorea
34		98.9	1262	238	7	ADF70380 Adf70380 Aequorea
35		98.9	1262	238	7	ADM78505 Adm78505 Wild-type
36		98.9	1262	238	7	ADM78577 Adm78577 Mutant Ae
37		98.9	1262	238	8	ADQ59552 Adq59552 Aequorea
38		98.9	1262	239	8	ADS17705 Ads17705 Green flu
39		98.9	1262	432	5	ABB08630 Abb08630 GFP fusio
40		98.9	1262	441	5	ABB08631 Abb08631 GFP fusio
41		98.9	1262	450	5	ABB08632 Abb08632 GFP fusio
42		98.9	1262	468	5	ABB08633 Abb08633 GFP fusio
43		98.9	1262	477	5	ABB08634 Abb08634 GFP fusio
44		98.9	1262	491	8	ADS17719 Ads17719 Green flu
45		98.9	1262	554	2	AAW48661 Aaw48661 RG fusion

```
                                  ALIGNMENTS

RESULT 1
ADA25196
ID   ADA25196 standard; protein; 238 AA.
XX
AC   ADA25196;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Green fluorescent protein mutant, F64L-S65T-S175G-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key            Location/Qualifiers
FT   Misc-difference 64
FT                  /note= "Leu replaces wild-type Phe"
FT   Active-site    65
FT                  /note= "Thr replaces wild-type Ser"
FT   Misc-difference 175
FT                  /note= "Gly replaces wild-type Ser"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-0023288.
XX
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AB, Michael NP,  Thomas N;
XX
DR   WPI; 2003-095652/09.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Claim 9; Fig 4; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid substitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proistion 65
```

PA   (AMSH) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AE, Michael NP,   Thomas N;
XX
DR   WPI; 2003-095652/09.
DR   N-PSDB; ADA25193.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Example 2; Page; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid substitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proistion 65
CC   or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), as
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial
CC   detection. The present sequence represents an Aequorea victoria GFP
CC   mutant used in an example of the invention. Note: The present sequence is
CC   not shown in the specification, but is derived from the wild-type GFP
CC   sequence shown in Fig 2 and the information given on page 24.
XX
SQ   Sequence 238 AA;

Query Match   99.7%; Score 1272; DB 6; Length 238;
Best Local Similarity   99.6%; Pred. No. 5.4e-125;
Matches 237; Conservative   1; Mismatches   0;   Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 3
AAW76106
ID   AAW76106 standard; protein; 238 AA.
XX
AC   AAW76106;
XX
DT   18-NOV-1998 (first entry)
XX
DE   A. victoria green fluorescent protein mutant F64L/S65T.
XX
KW   Green fluorescent protein; GFP; mutant; jellyfish; excitation;
KW   chromophore.
XX

CC   or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), as
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial
CC   detection. The present sequence represents the specifically claimed
CC   Aequorea victoria GFP mutant F64L-S65T-S175G-GFP.
XX
SQ   Sequence 238 AA;

Query Match   100.0%; Score 1276; DB 6; Length 238;
Best Local Similarity   100.0%; Pred. No. 2.1e-125;
Matches 238; Conservative   0; Mismatches   0;   Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 2
ADA25218
ID   ADA25218 standard; protein; 238 AA.
XX
AC   ADA25218;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Green fluorescent protein mutant, F64L-S175G-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 64
FT   /note= "Leu replaces wild-type Phe"
FT   Misc-difference 175
FT   /note= "Gly replaces wild-type Ser"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-00023288.
XX
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH) AMERSHAM PHARMACIA BIOTECH UK LTD.

us-10-757-624-4.rag

```
OS   Aequorea victoria.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 64
FT                   /label= F64L
FT                   /note= "Wild-type Phe is replaced with Leu"
FT   Misc-difference 65
FT                   /label= S65T
FT                   /note= "Wild-type Ser is replaced by Thr"
XX
PN   US5804387-A.
XX
PD   08-SEP-1998.
XX
PP   31-JAN-1997;  97US-00791332.
XX
PR   01-FEB-1996;  96US-0010960P.
XX
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI   Cormack BP, Valdivia RH, Falkow S;
XX
DR   WPI; 1998-505643/43.
XX
PT   DNA encoding mutant green fluorescent pigment proteins - with greater
PT   fluorescence intensity than wild-type proteins, useful for studying gene
PT   expression and protein localisation.
XX
PS   Disclosure; Page; 15pp; English.
XX
CC   This sequence represents a mutant green fluorescent protein (GFP) from
CC   Aequorea victoria in which a Phe residue at position 64 in the wild type
CC   protein is replaced by a Leu and a Ser at position 65 is replaced by a
CC   Thr. These mutations occur in the chromophore region. GFP mutants are
CC   used in a method to discover GFP's mutated in the chromophore region that
CC   fluoresce more brightly than wild-type GFP upon excitation at 488 nm.
CC   These mutants can be used in a method for analyzing a cell containing the
CC   GFP. The GFP can also be fused to a protein and used to identify the
CC   intracellular localisation of a protein of interest. A regulatory element
CC   could be operatively connected to a coding portion encoding a mutant GFP
CC   and exposed to an environmental stimulus. The fluorescence signal from
CC   the from the cell then measures the effect of the stimulus on the
CC   regulatory element. NOTE: This sequence does not appear in the
CC   specification but has been constructed from the wild-type GFP protein
CC   represented in AAW76106
XX
SQ   Sequence 238 AA;

Query Match         99.5%;  Score 1270;  DB 2;  Length 238;
Best Local Similarity 99.6%;  Pred. No. 8.8e-125;
Matches 237;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy        1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy       61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
            ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db       61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRABVKFEGDTLV 120

Qy      121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db      121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy      181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 4
ADA25214
IQ   ADA25214 standard; protein; 238 AA.
```

```
XX
AC   ADA25214;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Green fluorescent protein mutant, S175G-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 175
FT                   /note= "Gly replaces wild-type Ser"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PP   28-SEP-2001; 2001GB-0002328B.
XX
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AB, Michael NP, Thomas N;
XX
DR   WPI; 2003-095652/09.
DR   N-PSDB; ADA25193.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Example 2; Page; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid substitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proistion 65
CC   or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), as
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial
CC   detection. The present sequence represents an Aequorea victoria GFP
CC   mutant used in an example of the invention. Note: The present sequence is
CC   not shown in the specification, but is derived from the wild-type GFP
CC   sequence shown in Fig 2 and the information given on page 24.
XX
SQ   Sequence 238 AA;

Query Match         99.4%;  Score 1268;  DB 6;  Length 238;
Best Local Similarity 99.2%;  Pred. No. 1.4e-124;
Matches 236;  Conservative 1;  Mismatches 1;  Indels 0;  Gaps 0;

Qy        1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
```

us-10-757-624-4.rag

```
Qy    61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
          ||| :||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 5
ABR44424
ID   ABR44424 standard; protein; 238 AA.
XX
AC   ABR44424;
XX
DT   11-JUL-2003 (first entry)
XX
DE   F64L-Y66H-S175G-GFP #SEQ ID 3.
XX
KW   Green fluorescent protein; GFP; marker; label; reporter;
KW   bacterial detection; mutein.
XX
OS   Aequorea victoria.
XX
PN   WO200302286-A1.
XX
PD   10-APR-2003.
XX
PF   27-SEP-2002; 2002WO-GB004354.
XX
PR   28-SEP-2001; 2001GB-00023314.
XX
PA   (AMSH) AMERSHAM BIOSCIENCES UK LTD.
PA   (STUB/) STUBBS S.
XX
PI   Jones AE, Davies J, Ruehlmann D, Michael NP;
XX
DR   WPI; 2003-421211/39.
XX
PT   Novel blue-shifted variants of green fluorescent protein having improved
PT   fluorescence properties, useful for measuring expression of a protein of
PT   interest in a cell, as a secretion, cell or organelle marker.
XX
PS   Claim 7; Fig 3; 45pp; English.
XX
CC   The invention relates to a fluorescent protein derived from Green
CC   fluorescent Protein (GFP), or any functional GFP analog having an amino
CC   acid sequence which is modified by amino acid substitution compared with
CC   the sequence of wild type (wt) GFP. The polynucleotide encoding a GFP of
CC   the invention is useful for measuring the expression of a protein of
CC   interest in a cell, and is also useful for determining the cellular
CC   and/or extracellular localisation of a protein of interest. GFP's of the
CC   invention are useful as non-toxic markers for selection of transfected
CC   cells containing an expression vector encoding at least the fluorescent
CC   protein. GFP's may also be used as protein labels in living and fixed
CC   cells, as markers in cell or organelle fusion, for visualising
CC   translocation of proteins fused to them, as secretion markers, as genetic
CC   reporters or protein tags in transgenic animals, as cell or organelle
CC   activated cell sorting (FACS) and as reporters for bacterial detection.
CC   GFP's are also useful for performing transposon mutagenesis, where the
CC   GFP is used as a marker in transcriptional and translational fusions.
CC   GFP's of the invention have improved fluorescence properties and are
CC   suitable labels for proteins present at low concentrations. The current
CC   sequence represents the green fluorescent protein (GFP) mutant F64L-Y66H-
CC   S175G-GFP
XX
SQ   Sequence 238 AA;

Query Match   99.3%;   Score 1267;   DB 6;   Length 238;
Best Local Similarity   99.2%; Pred. No. 1.8e-124;
Matches 236;   Conservative 2; Mismatches 2;   Indels 0;   Gaps 0;

```
Qy    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60

Qy   61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         ||| :||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VTTLSHGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 6
AAW22101
ID   AAW22101 standard; protein; 238 AA.
XX
AC   AAW22101;
XX
DT   10-MAR-1998 (first entry)
XX
DE   Aequorea victoria protein GFP10 mutant F65L+S65T.
XX
KW   Green fluorescent protein; GFP; jellyfish; gfp10 gene; reporter;
KW   humanise; expression; mutant.
XX
OS   Aequorea victoria.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 64
FT   /label= F64L
FT   /note= "wild-type Phe is replaced by Leu"
FT   Misc-difference 65
FT   /label= S65T
FT   /note= "wild-type Ser is replaced by Thr"
XX
PN   WO9726333-A1.
XX
PD   24-JUL-1997.
XX
PF   17-JAN-1997; 97WO-US000755.
XX
PR   18-JAN-1996; 96US-00588201.
XX
PA   (UYFL) UNIV FLORIDA RES FOUND INC.
XX
PI   Zolotukhin S, Muzyczka N, Hauswirth WW;
XX
DR   WPI; 1997-385337/35.
XX
PT   Humanised green fluorescent protein gene - optimised to provide high
PT   level expression in mammalian cells, used e.g. to label or identify
PT   cells, to locate proteins, etc.
XX
PS   Claim 115; Page; 158pp; English.
XX
CC   This sequence represents a mutant form of the green fluorescent protein
CC   GFP10 i.e. F64L+S65T. Green fluorescent proteins (GFP's) could be used as
CC   reporter molecules as they absorbs blue light and emits green light
CC   without requiring any cofactors, substrates, or additional gene products
CC   allowing GFP detection in living cells providing meaningful gene
CC   expression in living cells. By providing humanised GFP, adapted for
CC   expression in mammalian and human cells, problems associated with wild-
CC   type jellyfish GFP e.g.variable and low expression levels should be
CC   overcome. Note: The present sequence does not appear in the
```

CC  sample contains protein kinase (PK) activity. The method involves
CC  contacting the sample with a phosphate donor (PD) and a fluorescent
CC  protein (FP) substrate for a PK, the protein substrate comprising a FP
CC  moiety and a phosphorylation site for a PK, where the protein substrate
CC  exhibits a different fluorescent property in the phosphorylated state
CC  than in the un-phosphorylated state. The protein substrate is then
CC  excited and the amount of a fluorescent property that differs in the un-
CC  phosphorylated state and phosphorylated state is measured, whereby an
CC  amount that is consistent with the presence of the protein substrate in
CC  its phosphorylated state indicates the presence of PK activity. The
CC  method and products can be used in drug screening. They can be used for
CC  screening for compounds which affect cellular events, including receptor-
CC  ligand binding, protein-protein interactions or kinase activation, which
CC  signal to the target kinase. NOTE: This sequence does not appear in the
CC  specification and has been constructed from the wild-type sequence
CC  represented in AAW40479
XX
SQ  Sequence 238 AA;

Query Match             99.2%;  Score 1266;  DB 2;  Length 238;
Best Local Similarity   99.2%;  Pred. No. 2.3e-124;
Matches  236;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy      61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRABVKFEGDTLV 120
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      61 VTTFTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRABVKFEGDTLV 120

Qy     121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 7
AAW65078
ID  AAW65078 standard; protein; 238 AA.
XX
AC  AAW65078;
XX
DT  15-SEP-1998  (first entry)
XX
DE  A. victoria green fluorescent protein variant S65T.
XX
KW  Green fluorescent protein; protein kinase; phosphate donor;
KW  phosphorylation; drug screening; receptor-ligand binding; signalling;
KW  protein-protein interaction; kinase activation.
XX
OS  Aequorea victoria.
OS  Synthetic.
XX
FH  Key    Location/Qualifiers
FT  Misc-difference   65
FT                    /label= S65T
FT                    /note= "wild type Ser is replaced with Thr"
XX
PN  WO9802571-A1.
XX
PD  22-JAN-1998.
XX
PF  16-JUL-1997;   97WO-US012410.
XX
PR  16-JUL-1996;   96US-00679865.
PR  16-JUL-1996;   96US-00680876.
PR  16-JUL-1996;   96US-00680877.
XX
PA  (REGC ) UNIV CALIFORNIA.
XX
PI  Tsien RY, Cubitt AB;
XX
DR  WPI; 1998-110616/10.
XX
PT  Assays for protein kinase and modulators - using a fluorescent protein
PT  substrate which exhibits a different fluorescent property in the
PT  phosphorylated and un-phosphorylated state.
XX
PS  Claim 27; Page; 65pp; English.
XX
CC  This sequence represents a green fluorescent protein variant from
CC  Aequorea victoria in which a wild-type Ser residue at position 65 is
CC  replaced with a Thr. This variant is used in a method to determine if a

---

CC  specification. It has been made by modifying the wild-type GFP10 amino
CC  acid sequence found in AAW31295
XX
SQ  Sequence 238 AA;

Query Match             99.2%;  Score 1266;  DB 2;  Length 238;
Best Local Similarity   99.2%;  Pred. No. 2.3e-124;
Matches  236;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;

Qy       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy      61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRABVKFEGDTLV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRABVKFEGDTLV 120

Qy     121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 8
AAW96328
ID  AAW96328 standard; protein; 238 AA.
XX
AC  AAW96328;
XX
DT  28-JUN-1999  (first entry)
XX
DE  Humanised green fluorescent protein.
XX
KW  Green fluorescent protein; gfp; jellyfish; Aequorea victoria;
KW  humanisation; reporter gene; substrate; cofactor; beta galactosidase;
KW  firefly luciferase; alkaline phosphatase;
KW  chloramphenicol acetyltransferase; CAT; beta glucuronidase; GUS.
XX
OS  Synthetic.
OS  Aequorea victoria.
XX
PN  WO9903997-A1.
XX
PD  28-JAN-1999.
XX
PF  16-JUL-1998;   98WO-US014692.
XX
PR  16-JUL-1997;   97US-00893327.
XX
PA  (UYFL ) UNIV FLORIDA.
XX
PI  Muzyczka N, Zolotukhin S,  Hauswirth W;
XX
DR  WPI; 1999-132241/11.
DR  N-PSDB; AAX08454.
XX

PT   Humanised green fluorescent protein - used to measure gene expression and
PT   identify transformed cells.
XX
PS   Disclosure; Page 136-137; 152pp; English.
XX
CC   Humanised green fluorescent protein (gfp) genes can be used to identify
CC   transformed cells, to measure gene expression in vitro and in vivo, to
CC   label specific cells in multicellular organisms (e.g. to study cell
CC   lineage's), to label and locate fusion proteins, and to study
CC   intracellular trafficking. Commonly used reporter genes include beta-
CC   galactosidase, firefly luciferase, alkaline phosphatase; chloramphenicol
CC   acetyltransferase (CAT), and beta glucuronidase (GUS). However, these
CC   have limitations in their use. Frequently, these reporter genes require
CC   the addition of a substrate and the size of certain proteins means that
CC   the expression of reporter fusion proteins can be difficult. The light
CC   stimulated GFP fluorescence is species independent and does not require
CC   any cofactors substrates or additional gene products from Aequorea
CC   victoria an as the GFP genes have been humanised, they are expressed at
CC   sufficient levels to be detectable in human cells, unlike previous GFP
CC   proteins
XX
SQ   Sequence 238 AA;

Query Match    99.2%;   Score 1266;   DB 2;   Length 238;
Best Local Similarity   99.2%;   Pred. No. 2.3e-124;
Matches 236;   Conservative    1;   Mismatches  1;   Indels  0;   Gaps  0;

```
Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db   61 VTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 9
AAB16042
ID   AAB16042 standard; protein; 238 AA.
XX
AC   AAB16042;
XX
DT   26-MAR-2002   (first entry)
XX
DE   Aequorea victoria GFP mutant protein (S65T).
XX
KW   Green fluorescent protein; cell lineage tracer; protein localisation;
KW   GFP; fusion tag; gene expression marker; fluorescent energy acceptor;
KW   immunoassay; hybridisation assay; fluorescent energy donor; biosensor;
KW   FRET; fluorescence resonance energy transfer; mutant; mutein.
XX
OS   Aequorea victoria.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 65
FT   /note= "Wild type Ser substituted with Thr"
XX
PN   WO200190147-A2.
XX
PD   29-NOV-2001.
XX
PF   17-MAY-2001; 2001WO-US016149.
XX
PR   19-MAY-2000; 2000US-00575847.

XX
PA   (UYOR-) UNIV OREGON STATE.
XX
PI   Wachter R, Remington SJ;
XX
DR   WPI; 2002-083084/11.
XX
PT   New long wavelength engineered fluorescent proteins, useful as markers
PT   for gene expression, tracers of cell lineage or as fusion tags to monitor
PT   protein localization, or in detection assays, e.g. immunoassays or
PT   hybridization assays.
XX
PS   Claim 2; Page; 181pp; English.
XX
CC   The patent discloses functional engineered fluorescent protein and its
CC   corresponding polynucleotide. The amino acid sequence of the engineered
CC   protein is identical to Aequorea green fluorescent protein (GFP). The
CC   engineered fluorescent proteins of the invention have varied fluorescent
CC   properties and have the ability to respond to ion concentrations via a
CC   change in fluorescence characteristics. They are useful as markers for
CC   gene expression, tracers of cell lineage or as fusion tags to monitor
CC   protein localisation within living cells. The engineered fluorescent
CC   proteins are particularly useful for coupling engineered fluorescent
CC   proteins to antibodies, nucleic acids or other receptors for use in
CC   detection assays, e.g. immunoassays or hybridisation assays. They are
CC   useful for tracking the movement of proteins in cells or in systems for
CC   detecting induction of transcription and for the simultaneous measurement
CC   of two or more processes within cells. Proteins of the invention are also
CC   useful as fluorescent energy donors or acceptors as well as biosensors
CC   for detecting anions. They are also useful in fluorescence resonance
CC   energy transfer (FRET). The crystal structure of the GFP is useful for
CC   designing mutants having altered fluorescent characteristics which are
CC   particularly useful to identify amino acids whose substitution alters
CC   fluorescent properties of the protein. The crystal structure of the GFP
CC   is also useful for designing mutants having altered anion binding
CC   characteristics which are particularly useful for identifying amino acids
CC   whose substitution alters the specificity and affinity of the binding
CC   site to various anions, and for monitoring anion binding and therefore
CC   the concentration of the anion. The present sequence is Aequorea victoria
CC   GFP mutant protein (S65T). Note: This sequence is not shown in the
CC   specification but is derived from Aequorea victoria wild type green
CC   fluorescent protein shown in figure 3 of the specification (AAB16038)
XX
SQ   Sequence 238 AA;

Query Match   99.2%;   Score 1266;   DB 5;   Length 238;
Best Local Similarity   99.2%;   Pred. No. 2.3e-124;
Matches 236;   Conservative   0;   Mismatches 2;   Indels 0;   Gaps 0;

```
Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        ||||:|||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 10
ABG76008
ID   ABG76008 standard; protein; 238 AA.
XX
AC   ABG76008;
XX
DT   30-APR-2003   (first entry)

```
Db   61 VTTFTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKPEGDTLV 120
        |||  |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNPKIRHNIEDGSVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||  ||||

Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNPKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 11
ADA25212
ID   ADA25212 standard; protein; 238 AA.
XX
AC   ADA25212;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Green fluorescent protein mutant, F64L-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 64
FT   /note= "Leu replaces wild-type Phe"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-0023288.
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AE, Michael NP, Thomas N;
XX
DR   WPI; 2003-095652/09.
DR   N-PSDB; ADA25193.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Example 2; Page; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid substitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proteion 65
CC   or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), as
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial

XX
DE   Jellyfish GFP mutant S65T.
XX
KW   Jellyfish; enzyme; green fluorescent protein; GFP; FRET; mutant;
KW   fluorescent resonance energy transfer; tandem fluorescent protein;
KW   enzymatic assay; Alzheimer's disease; hypertension; inflammation;
KW   apoptosis; AIDS; acquired immunodeficiency syndrome; mutein.
XX
OS   Aequorea victoria.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 65
FT   /note= "Wild-type Ser substituted by Thr"
XX
PN   US2002164674-A1.
XX
PD   07-NOV-2002.
XX
PF   25-JAN-2002; 2002US-00057505.
XX
PR   31-JAN-1996; 96US-00594575.
PR   31-JAN-1997; 97US-00792553.
PR   13-SEP-1999; 99US-00396003.
XX
PA   (REGC) UNIV CALIFORNIA.
XX
PI   Tsien RY, Heim R, Cubitt A;
XX
DR   WPI; 2003-247255/24.
XX
PT   New tandem fluorescent protein construct comprising a donor or acceptor
PT   fluorescent protein moiety or a linker moiety that couples the donor and
PT   acceptor moieties, useful in enzymatic assays.
XX
PS   Claim 3; Page; 34pp; English.
XX
CC   The invention relates to a tandem fluorescent protein construct
CC   comprising a donor or acceptor fluorescent protein moiety (e.g. jellyfish
CC   green fluorescent protein, GFP) or a linker moiety that couples the donor
CC   and acceptor moieties. Also include are a recombinant nucleic acid coding
CC   for expression of the tandem fluorescent protein construct, an expression
CC   vector comprising expression control sequences operatively linked to a
CC   sequence coding for the expression of the tandem fluorescent protein
CC   construct, a host cell transfected with the expression vector,
CC   determining whether a sample contains an enzyme or whether a compound
CC   alters the activity of an enzyme, determining the amount of activity of
CC   an enzyme in a cell and testing for cleavage activity. The tandem
CC   fluorescent protein construct is useful in enzymatic assays, using the
CC   principle of fluorescent resonance energy transfer (FRET) between the
CC   donor and acceptor moieties. The tandem fluorescent proteins are
CC   particularly useful in assays for protease activity. Proteases play an
CC   essential role in many disease processes e.g. Alzheimer's disease,
CC   hypertension, inflammation, apoptosis and AIDS (acquired immunodeficiency
CC   syndrome). Tandem fluorescent proteins were constructed comprising
CC   mutants of GFP with altered fluorescent spectra. The present sequence
CC   represents a mutant GFP suitable for incorporation into a tandem
CC   fluorescent protein of the invention. Note: The present sequence is not
CC   shown in the specification but was created by the indexer using the
CC   information in the specification and the wild-type GFP protein appearing
CC   as ABG75980

XX
SQ   Sequence 238 AA;

Query Match   99.2%;   Score 1266;   DB 6;   Length 238;
Best Local Similarity 99.2%;   Pred. No. 2.3e-124;
Matches 236;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 MSKGEELFTGVVPIIVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1 MSKGEELFTGVVPIIVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRABVKPEGDTLV 120
```

CC detection. The present sequence represents an Aequorea victoria GFP
CC mutant used in an example of the invention. Note: The present sequence is
CC not shown in the specification, but is derived from the wild-type GFP
CC sequence shown in Fig 2 and the information given on page 24.
XX
SQ Sequence 238 AA;

Query Match   99.2%;   Score 1266; DB 6; Length 238;
Best Local Similarity   99.2%;   Pred. No. 2.3e-124;
Matches   236; Conservative   1; Mismatches   1; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 12
ADA25224
ID ADA25224 standard; protein; 238 AA.
XX
AC ADA25224;
XX
DT 20-NOV-2003 (first entry)
XX
DE Green fluorescent protein mutant, F64L-S65T-V163A-GFP.
XX
KW Green fluorescent protein; GFP; jellyfish; marker protein;
KW reporter protein; mutant; mutein.
XX
OS Synthetic.
OS Aequorea victoria.
XX
FH Key   Location/Qualifiers
FT Misc-difference 64
FT   /note= "Leu replaces wild-type Phe"
FT Misc-difference 65
FT   /note= "Thr replaces wild-type Ser"
FT Misc-difference 163
FT   /note= "Ala replaces wild-type Val"
XX
GB2374868-A.
XX
PD 30-OCT-2002.
XX
PF 28-SEP-2001; 2001GB-00023288.
XX
PR 23-APR-2001; 2001GB-00009858.
XX
PA (AMSH) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA (AMSH) AMERSHAM BIOSCIENCES UK LTD.
XX
PI Stubbs SLJ, Jones AE, Michael NP, Thomas N;
XX
DR WPI; 2003-095652/09.
DR N-PSDB; ADA25193.
XX
PT Novel fluorescent protein derived from green fluorescent protein useful
PT as a transfection marker, has different excitation spectrum and/or
PT emission spectrum compared with wild-type green fluorescent protein.
XX
PS Example 2; Page; 52pp; English.

CC The invention relates to Aequorea victoria green fluorescent protein
CC (GFP) mutants containing an amino acid substitution at positions 64 and
CC 175, and additionally an amino acid substitution at either proistion 65
CC or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC invention exhibit enhanced fluorescence relative to wild type GFP when
CC expressed in non-homologous cells at temperatures above 30 degrees
CC Celsius, and excited at 490 nm. The mutants can also be detected in
CC mammalian cells at lower levels of expression and with increased
CC sensitivity relative to wild type GFP. The GFP mutants of the invention
CC are useful as non-toxic markers for selection of transfected cells, as
CC protein labels in living and fixed cells, as markers in cell or organelle
CC fusion, for visualising translocation of intracellular proteins to a
CC specific organelle, as secretion markers, as genetic reporters or protein
CC tags for protein and gene expression in transgenic animals, as cell or
CC organelle integrity markers, as transfection markers, as markers to be
CC used in combination with fluorescent activated cell sorting (FACS), as
CC real-time probes working at near physiological concentrations, for
CC performing transposon vector mutagenesis, and as reporters for bacterial
CC detection. The present sequence represents an Aequorea victoria GFP
CC mutant used in an example of the invention. Note: The present sequence is
CC not shown in the specification, but is derived from the wild-type GFP
CC sequence shown in Fig 2 and the information given on page 24.
XX
SQ Sequence 238 AA;

Query Match   99.2%;   Score 1266; DB 6; Length 238;
Best Local Similarity   99.2%;   Pred. No. 2.3e-124;
Matches   236; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKIRHNIEDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 13
AAW96329
ID AAW96329 standard; protein; 247 AA.
XX
AC AAW96329;
XX
DT 28-JUN-1999 (first entry)
XX
DE Humanised green fluorescent protein.
XX
KW Green fluorescent protein; gfp; jellyfish; Aequorea victoria;
KW humanisation; reporter gene; substrate; cofactor; beta galactosidase;
KW firefly luciferase; alkaline phosphatae;
KW chloramphenicol acetyltransferase; CAT; beta glucuronidase; GUS.
XX
OS Synthetic.
OS Aequorea victoria.
XX
PN WO903997-A1.
XX
PD 28-JAN-1999.
XX
PF 16-JUL-1998; 98WO-US014692.
XX
PR 16-JUL-1997; 97US-00893327.

```
FT   Misc-difference  65
FT                    /label= S65C
FT                    /note= "wild type Ser substituted with Thr"
FT   Peptide          239..264
FT                    /label= linker_moeity
FT   Cleavage-site    250..251
FT                    /label= trypsin_cleavage_site
FT   Cleavage-site    253..254
FT                    /label= calpain_cleavage_site
FT   Cleavage-site    258..259
FT                    /label= trypsin_enterokinase_cleavage_site
FT   Protein          265..501
FT                    /label= W7_GFP_variant
FT   Misc-difference  329
FT                    /label= Y66W
FT                    /note= "wild type Tyr substituted with Trp"
FT   Misc-difference  409
FT                    /label= N146L
FT                    /note= "wild type Asn substituted with Leu"
FT   Misc-difference  416
FT                    /label= M153T
FT                    /note= "wild type Met substituted with Thr"
FT   Misc-difference  426
FT                    /label= V163A
FT                    /note= "wild type Val substituted with Ala"
FT   Misc-difference  475
FT                    /label= N212K
FT                    /note= "wild type Asn substituted with Lys"
XX
PN   WO9728261-A1.
XX
PD   07-AUG-1997.
XX
PF   31-JAN-1997;   97WO-US001457.
XX
PR   31-JAN-1996;   96US-00594575.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Tsien RY, Heim R, Cubitt A;
XX
DR   WPI; 1997-402615/37.
XX
PT   Tandem fluorescent protein constructs - have donor and acceptor moieties
PT   exhibiting fluorescent linked via cleavable peptide linker, useful in
PT   enzymatic assays.
XX
PS   Claim 3; Page; 88pp; English.
XX
CC   This protein sequence is that of a novel tandem fluorescent protein
CC   construct, made using Aequorea victoria (North West Pacific jellyfish)
CC   green fluorescent protein (GFP) variants S65T and W7. W7 fluoresces at a
CC   shorter wavelength than GFP. The construct comprises a donor (e.g. S65T)
CC   and an acceptor (e.g. W7) fluorescent protein moiety (donors and
CC   acceptors can be green or blue fluorescent proteins), and a linker
CC   coupling them. Preferably, the donor is positioned at the N-terminus of
CC   the polypeptide relative to the acceptor. The linker moiety is a peptide
CC   5-50 amino acids in length containing a protease cleavage site. In this
CC   example, the linker moiety contains many recognition sites for proteases,
CC   including trypsin, calpain and enterokinase. The donor and acceptor
CC   moieties exhibit fluorescent resonance energy transfer (FRET) when they
CC   are cleaved. The constructs are used in enzymatic assays and can be used
CC   to isolate new enzymes or enzyme inhibitors or promoters. The specific
CC   activity of enzyme (in vivo and in vitro) and compounds altering enzyme
CC   activity can be obtained. FRET and hence activity of specific compounds
CC   is measured from the acceptor or donor moiety or maybe obtained using a
CC   ratio between the two. Note: The present sequence does not appear in the
CC   specification; it has been made by modifying the native GFP sequence, and
CC   adding the linker moiety in the appropriate place
XX
SQ   Sequence 501 AA;
```

```
XX
PA   (UYFL ) UNIV FLORIDA.
XX
PI   Muzyczka N, Zolotukhin S, Hauswirth W;
XX
DR   WPI; 1999-132241/11.
DR   N-PSDB; AAX08455.
XX
PT   Humanised green fluorescent protein - used to measure gene expression and
PT   identify transformed cells.
XX
PS   Claim 5; Page 142; 152pp; English.
XX
CC   Humanised green fluorescent protein (gfp) genes can be used to identify
CC   transformed cells, to measure gene expression in vitro and in vivo, to
CC   label specific cells in multicellular organisms (e.g. to study cell
CC   lineage's), to label and locate fusion proteins, and to study
CC   intracellular trafficking. Commonly used reporter genes include beta-
CC   galactosidase, firefly luciferase, alkaline phosphatase; chloramphenicol
CC   acetyltransferase (CAT), and beta glucuronidase (GUS). However, these
CC   have limitations in their use. Frequently, these reporter genes require
CC   the addition of a substrate and the size of certain proteins means that
CC   the expression of reporter fusion proteins can be difficult. The light
CC   stimulated GFP fluorescence is species independent and does not require
CC   any cofactors substrates or additional gene products from Aequorea
CC   victoria an as the GFP genes have been humanised, they are expressed at
CC   sufficient levels to be detectable in human cells, unlike previous GFP
CC   proteins
XX
SQ   Sequence 247 AA;

Query Match             99.2%;  Score 1266;  DB 2;  Length 247;
Best Local Similarity   99.2%;
Matches  236;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  69

Qy   61 VTTLTTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         ||||||||||||||||||||||| ::|||||||||||||||||||||||||||||||||||
Db   70 VTTLTTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 129

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  130 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 189

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  190 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 247

RESULT 14
AAW31879
ID   AAW31879 standard; protein; 501 AA.
XX
AC   AAW31879;
XX
DT   03-FEB-1998   (first entry)
XX
DE   GFP variants S65T and W7 tandem fluorescent protein construct.
XX
KW   North West Pacific jellyfish; green fluorescent protein; GFP; S65C;
KW   fluorescent resonance energy transfer; FRET; enzymatic assay; W7;
KW   enzyme inhibitor; enzyme promoter; PCR primer; protease cleavage site;
KW   Tandem fluorescent protein construct; blue fluorescent protein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key          Location/Qualifiers
FT   Protein      1..238
FT                /label= S65T_GFP_variant
```

```
Query Match          99.2%;    Score 1266;  DB 2;   Length 501;
Best Local Similarity  99.2%;  Pred. No. 7e-124;
Matches  236;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

```
RESULT 15
ADA25195
ID   ADA25195 standard; protein; 238 AA.
XX
AC   ADA25195;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Green fluorescent protein mutant, F64L-S175G-E222G-GFP.
XX
KW   Green fluorescent protein; GFP; jellyfish; marker protein;
KW   reporter protein; mutant; mutein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
FH   Key      Location/Qualifiers
FT   Misc-difference  64
FT              /note= "Leu replaces wild-type Phe"
FT   Misc-difference  175
FT              /note= "Gly replaces wild-type Ser"
FT   Misc-difference  222
FT              /note= "Gly replaces wild-type Glu"
XX
PN   GB2374868-A.
XX
PD   30-OCT-2002.
XX
PF   28-SEP-2001; 2001GB-00023288.
PR   23-APR-2001; 2001GB-00009858.
XX
PA   (AMSH ) AMERSHAM PHARMACIA BIOTECH UK LTD.
PA   (AMSH ) AMERSHAM BIOSCIENCES UK LTD.
XX
PI   Stubbs SLJ, Jones AE, Michael NP, Thomas N;
XX
DR   WPI; 2003-095652/09.
XX
PT   Novel fluorescent protein derived from green fluorescent protein useful
PT   as a transfection marker, has different excitation spectrum and/or
PT   emission spectrum compared with wild-type green fluorescent protein.
XX
PS   Claim 8; Fig 3; 52pp; English.
XX
CC   The invention relates to Aequorea victoria green fluorescent protein
CC   (GFP) mutants containing an amino acid substitution at positions 64 and
CC   175, and additionally an amino acid substitution at either proistion 65
CC   or 222. The mutants of the invention are particularly F64L-S175G-E222G-
CC   GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
CC   invention exhibit enhanced fluorescence relative to wild type GFP when
CC   expressed in non-homologous cells at temperatures above 30 degrees
CC   Celsius, and excited at 490 nm. The mutants can also be detected in
CC   mammalian cells at lower levels of expression and with increased
CC   sensitivity relative to wild type GFP. The GFP mutants of the invention
CC   are useful as non-toxic markers for selection of transfected cells, as
CC   protein labels in living and fixed cells, as markers in cell or organelle
CC   fusion, for visualising translocation of intracellular proteins to a
CC   specific organelle, as secretion markers, as genetic reporters or protein
CC   tags for protein and gene expression in transgenic animals, as cell or
CC   organelle integrity markers, as transfection markers, as markers to be
CC   used in combination with fluorescent activated cell sorting (FACS), as
CC   real-time probes working at near physiological concentrations, for
CC   performing transposon vector mutagenesis, and as reporters for bacterial
CC   detection. The present sequence represents the specifically claimed
CC   Aequorea victoria GFP mutant F64L-S175G-E222G-GFP.
XX
SQ   Sequence 238 AA;
```

```
Query Match          99.1%;    Score 1265;  DB 6;   Length 238;
Best Local Similarity  99.2%;  Pred. No. 3e-124;
Matches  236;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
```

Search completed: September 16, 2005, 17:16:11
Job time : 118 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2005, 17:08:27 ; Search time 31 Seconds
                                            (without alignments)
                                            573,112 Million cell updates/sec

Title: US-10-757-624-4
Perfect score: 1276
Sequence: 1 MSKGEELFTGVVPILVELDG..........VLLEFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
               1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
               2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
               3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
               4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
               5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep:*
               6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1266	99.2	238	3	US-08-893-327-16	Sequence 16, Appl
2	1266	99.2	247	3	US-08-893-327-18	Sequence 18, Appl
3	1262	98.9	238	1	US-08-753-143-2	Sequence 2, Appl
4	1262	98.9	238	2	US-08-679-865-2	Sequence 2, Appli
5	1262	98.9	238	2	US-08-680-876-2	Sequence 2, Appli
6	1262	98.9	238	3	US-08-792-553-2	Sequence 2, Appli
7	1262	98.9	238	3	US-08-753-144-2	Sequence 2, Appli
8	1262	98.9	238	3	US-09-094-359-2	Sequence 2, Appli
9	1262	98.9	238	3	US-09-172-063-2	Sequence 2, Appli
10	1262	98.9	238	4	US-09-263-975-2	Sequence 2, Appli
11	1262	98.9	238	4	US-08-727-452-2	Sequence 2, Appli
12	1262	98.9	238	4	US-09-418-785-1	Sequence 1, Appli
13	1262	98.9	238	4	US-09-129-192C-2	Sequence 2, Appli
14	1262	98.9	238	4	US-09-129-192C-74	Sequence 74, Appl
15	1262	98.9	238	4	US-09-602-641-2	Sequence 2, Appli
16	1262	98.9	238	4	US-09-704-463-2	Sequence 2, Appli
17	1262	98.9	238	4	US-09-472-065A-2	Sequence 2, Appli
18	1262	98.9	238	4	US-10-024-686A-2	Sequence 2, Appli
19	1262	98.9	238	4	US-08-594-575C-2	Sequence 2, Appli
20	1262	98.9	432	4	US-09-863-901-1	Sequence 1, Appli
21	1262	98.9	441	4	US-09-863-901-2	Sequence 2, Appli
22	1262	98.9	450	4	US-09-863-901-3	Sequence 3, Appli
23	1262	98.9	468	4	US-09-863-901-4	Sequence 4, Appli
24	1262	98.9	477	4	US-09-863-901-5	Sequence 5, Appli
25	1262	98.9	906	4	US-09-863-901-6	Sequence 6, Appli
26	1262	98.9	1070	4	US-09-091-042A-2	Sequence 2, Appli
27	1262	98.9	1452	3	US-09-127-227-2	Sequence 2, Appli
28	1259	98.7	238	4	US-09-472-065A-6	Sequence 6, Appli
29	1258	98.6	238	1	US-08-337-915A-2	Sequence 2, Appli
30	1258	98.6	238	3	US-09-121-539-1	Sequence 1, Appli
31	1258	98.6	238	4	US-09-214-909-2	Sequence 2, Appli
32	1258	98.6	238	4	US-09-479-645A-10	Sequence 10, Appl
33	1258	98.6	238	4	US-09-479-645A-159	Sequence 159, APP
34	1258	98.6	238	4	US-09-472-065A-4	Sequence 4, Appli
35	1258	98.6	238	4	US-09-920-922-4	Sequence 4, Appli
36	1258	98.6	238	5	PCT-US95-14692-2	Sequence 35, Appli
37	1257	98.5	238	3	US-09-023-946B-35	Sequence 2, Appli
38	1257	98.5	239	3	US-08-646-538-2	Sequence 5, Appli
39	1257	98.5	238	3	US-09-503-222-2	Sequence 2, Appli
40	1256	98.4	238	4	US-09-472-065A-5	Sequence 94, Appl
41	1256	98.4	238	4	US-09-989-025A-2	Sequence 96, Appl
42	1255	98.4	243	4	US-09-479-645A-94	Sequence 98, Appl
43	1255	98.4	243	4	US-09-479-645A-96	Sequence 100, App
44	1255	98.4	243	4	US-09-479-645A-98	
45	1255	98.4	243	4	US-09-479-645A-100	

ALIGNMENTS

RESULT 1
US-08-893-327-16
; Sequence 16, Application US/08893327
; Patent No. 6020192
; GENERAL INFORMATION:
;   APPLICANT: Zolotukhin, Sergei
;   APPLICANT: Hauswirth, William W.
;   APPLICANT: Muzyczka, Nicholas
;   TITLE OF INVENTION: Humanized Green Fluorescent Protein
;   TITLE OF INVENTION: Genes and Methods
;   NUMBER OF SEQUENCES: 20
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Arnold, White & Durkee
;   STREET: P. O. Box 4433
;   CITY: Houston
;   STATE: TX
;   COUNTRY: USA
;   ZIP: 77210-4433
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/893,327
;   FILING DATE:
;   CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/588,201
;   FILING DATE: 18-JAN-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Kitchell, Barbara S.
;   REGISTRATION NUMBER: 33,928
;   REFERENCE/DOCKET NUMBER: UFLA:062\KIT
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (512) 418-3000
;   TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 16:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 238 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-893-327-16

Query Match   99.2%; Score 1266; DB 3; Length 238;
Best Local Similarity   99.2%; Pred. No. 2.8e-130;
Matches 236; Conservative   1; Mismatches   1; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

```
Db           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
Db           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db           |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTTFTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Qy    121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
Db           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy    121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSSVQLAD 180
Qy    181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
Db           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 2
US-08-893-327-18
; Sequence 18, Application US/08893327
Patent No. 6020192
; GENERAL INFORMATION:
; APPLICANT: Zolotukhin, Sergei
; APPLICANT: Hauswirth, William W.
; APPLICANT: Muzyczka, Nicholas
; TITLE OF INVENTION: Humanized Green Fluorescent Protein
; TITLE OF INVENTION: Genes and Methods
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Arnold, White & Durkee
;   STREET: P. O. Box 4433
;   CITY: Houston
;   STATE: TX
;   COUNTRY: USA
;   ZIP: 77210-4433
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/893,327
;   FILING DATE:
;   CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/588,201
;   FILING DATE: 18-JAN-1996
; ATTORNEY/AGENT INFORMATION:
;   NAME: Kitchell, Barbara S.
;   REGISTRATION NUMBER: 33,928
;   REFERENCE/DOCKET NUMBER: UFLA:062\KIT
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (512) 418-3000
;   TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 18:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 247 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-893-327-18

Query Match   99.2%; Score 1266; DB 3; Length 247;
Best Local Similarity   99.2%; Pred. No. 3e-130;
Matches 236; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
Db           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  69
Qy     61 VTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db           |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     70 VTTFTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 129
```

RESULT 3
US-08-753-143-2
; Sequence 2, Application US/08753143A
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Heim, Roger
; TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
; FILE REFERENCE: 07257/032003
; CURRENT APPLICATION NUMBER: US/08/753,143A
; CURRENT FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Aequorea victoria
US-08-753-143-2

Query Match   98.9%; Score 1262; DB 1; Length 238;
Best Local Similarity   98.7%; Pred. No. 7.8e-130;
Matches 235; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
Db           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
Qy     61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db           |||::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Qy    121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
Db           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSSVQLAD 180
Qy    181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
Db           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 4
US-08-679-865-2
; Sequence 2, Application US/08679865
; Patent No. 5912137
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Assays for Protein Kinases Using
; TITLE OF INVENTION: Fluorescent Protein Substrates
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Townsend and Townsend and Crew LLP
;   STREET: Two Embarcadero Center, Eighth Floor
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94111-3834
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/679,865
;   FILING DATE: 16-JUL-1996

```
;         CLASSIFICATION:   435
;     ATTORNEY/AGENT INFORMATION:
;         NAME:  Storella, John S.
;         REGISTRATION NUMBER:  32,944
;         REFERENCE/DOCKET NUMBER:  0230TZ-069000
;     TELECOMMUNICATION INFORMATION:
;         TELEPHONE:   (415) 576-0200
;         TELEFAX:     (415) 576-0300
;  INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;         LENGTH:  238 amino acids
;         TYPE:  amino acid
;         TOPOLOGY:  linear
;         MOLECULE TYPE:  protein
US-08-679-865-2

Query Match             98.9%;  Score 1262;  DB 2;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 7.8e-130;
Matches  235;  Conservative   1;  Mismatches   2;  Indels    0;  Gaps    0;

Qy      1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy     61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
          ||| :||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy    121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy    181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 5
US-08-680-876-2
;  Sequence 2, Application US/08680876
;  Patent No. 5925558
;  GENERAL INFORMATION:
;     APPLICANT:  Tsien, Roger Y.
;     APPLICANT:  Cubitt, Andrew B.
;     TITLE OF INVENTION:  Assays for Protein Kinases Using
;     TITLE OF INVENTION:  Fluorescent Protein Substrates
;     NUMBER OF SEQUENCES:  48
;     CORRESPONDENCE ADDRESS:
;         ADDRESSEE:  Townsend and Townsend and Crew LLP
;         STREET:  Two Embarcadero Center, Eighth Floor
;         CITY:  San Francisco
;         STATE:  California
;         COUNTRY:  USA
;         ZIP:  94111-3834
;     COMPUTER READABLE FORM:
;         MEDIUM TYPE:  Floppy disk
;         COMPUTER:  IBM PC compatible
;         OPERATING SYSTEM:  PC-DOS/MS-DOS
;         SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;         APPLICATION NUMBER:  US/08/680,876
;         FILING DATE:  16-JUL-1996
;         CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;         NAME:  Storella, John S.
;         REGISTRATION NUMBER:  32,944
;         REFERENCE/DOCKET NUMBER:  0230TZ-069200
;     TELECOMMUNICATION INFORMATION:
;         TELEPHONE:   (415) 576-0200
;         TELEFAX:     (415) 576-0300
;  INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;         LENGTH:  238 amino acids
```

```
;         TYPE:  amino acid
;         TOPOLOGY:  linear
;         MOLECULE TYPE:  protein
US-08-680-876-2

Query Match             98.9%;  Score 1262;  DB 2;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 7.8e-130;
Matches  235;  Conservative   1;  Mismatches   2;  Indels    0;  Gaps    0;

Qy      1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy     61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
          ||| :||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy    121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy    181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 6
US-08-792-553-2
;  Sequence 2, Application US/08792553
;  Patent No. 5981200
;  GENERAL INFORMATION:
;     APPLICANT:  Tsien, Roger Y.
;     APPLICANT:  Helm, Roger
;     TITLE OF INVENTION:  Tandem Fluorescent Protein Constructs
;     NUMBER OF SEQUENCES:  25
;     CORRESPONDENCE ADDRESS:
;         ADDRESSEE:  FISH & RICHARDSON P.C.
;         STREET:  4225 Executive Square, Suite 1400
;         CITY:  La Jolla
;         STATE:  California
;         COUNTRY:  USA
;         ZIP:  92037
;     COMPUTER READABLE FORM:
;         MEDIUM TYPE:  Floppy disk
;         COMPUTER:  IBM PC compatible
;         OPERATING SYSTEM:  PC-DOS/MS-DOS
;         SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;         APPLICATION NUMBER:  US/08/792,553
;         FILING DATE:  31-JAN-1997
;         CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;         NAME:  Lisa A. Haile, Ph.D.:
;         REGISTRATION NUMBER:  38,347
;         REFERENCE/DOCKET NUMBER:  07257/041001/UC 96-160-2
;     TELECOMMUNICATION INFORMATION:
;         TELEPHONE:   619-678-5070
;         TELEFAX:     619-678-5099
;  INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;         LENGTH:  238 amino acids
;         TYPE:  amino acid
;         TOPOLOGY:  linear
;         MOLECULE TYPE:  protein
US-08-792-553-2

Query Match             98.9%;  Score 1262;  DB 2;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 7.8e-130;
Matches  235;  Conservative   1;  Mismatches   2;  Indels    0;  Gaps    0;

Qy      1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
```

```
Qy    61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
          |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 7
US-08-753-144-2
; Sequence 2, Application US/08753144
; Patent No. 6066476
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Heim, Roger
; TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windiws95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/753,144
; FILING DATE: 20-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/727,452
; FILING DATE: 10-OCT-1996
; APPLICATION NUMBER: US95/14692
; FILING DATE: 13-NOV-1995
; APPLICATION NUMBER: 08/337,915
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/032002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-753-144-2

Query Match   98.9%;   Score 1262;   DB 3;   Length 238;
Best Local Similarity   98.7%;   Pred. No. 7.8e-130;
Matches 235;   Conservative 1;   Mismatches 2;   Indels 0;   Gaps 0;

```
Qy    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
```

```
Qy   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 8
US-09-094-359-2
; Sequence 2, Application US/09094359
; Patent No. 6140132
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/067001
; CURRENT APPLICATION NUMBER: US/09/094,359
; CURRENT FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-094-359-2

Query Match   98.9%;   Score 1262;   DB 3;   Length 238;
Best Local Similarity   98.7%;   Pred. No. 7.8e-130;
Matches 235;   Conservative 1;   Mismatches 2;   Indels 0;   Gaps 0;

```
Qy    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 9
US-09-172-063-2
; Sequence 2, Application US/09172063
; Patent No. 6150176
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/172,063
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 09/094,359
; EARLIER FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT

; ORGANISM: Aequorea victoria
US-09-172-063-2

Query Match   98.9%;   Score 1262; DB 3; Length 238;
Best Local Similarity   98.7%; Pred. No. 7.8e-130;
Matches 235; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

```
Qy   1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy  61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db 121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 10
US-09-263-975-2
; Sequence 2, Application US/09263975
; Patent No. 6248550
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Assays for Protein Kinases Using
; TITLE OF INVENTION: Fluorescent Protein Substrates
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,975
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/679,865
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John S.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 02307Z-069000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-263-975-2

Query Match   98.9%;   Score 1262; DB 3; Length 238;
Best Local Similarity   98.7%; Pred. No. 7.8e-130;
Matches 235; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

```
Qy   1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy  61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db 121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 11
US-08-727-452-2
; Sequence 2, Application US/08727452A
; Patent No. 6319669
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Heim, Roger
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR DETECTION OF ANALYTES
; FILE REFERENCE: 07257/032001
; CURRENT APPLICATION NUMBER: US/08/727,452A
; CURRENT FILING DATE: 1996-03-20
; EARLIER APPLICATION NUMBER: PCT/US95/14692
; EARLIER FILING DATE: 1995-11-13
; EARLIER APPLICATION NUMBER: US 07/337,915
; EARLIER FILING DATE: 1994-11-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-08-727-452-2

Query Match   98.9%;   Score 1262; DB 3; Length 238;
Best Local Similarity   98.7%; Pred. No. 7.8e-130; Indels   0; Gaps   0;
Matches 235; Conservative   1; Mismatches   2;

```
Qy   1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy  61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db 121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy 181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 12
US-09-418-785-1
; Sequence 1, Application US/09418785
; Patent No. 6414119
; GENERAL INFORMATION:
; APPLICANT: Fisher, Hugh
; TITLE OF INVENTION: Rapidly Greening, Low Oxygen Mutant of
; TITLE OF INVENTION: the Aequorea victoria Green Fluorescent Protein
; FILE REFERENCE: RutC 99-0011
; CURRENT APPLICATION NUMBER: US/09/418,785
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/104,563
; PRIOR FILING DATE: 1998-10-16

```
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
; PUBLICATION INFORMATION:
; AUTHORS: Prasher, D.C. et al.
; TITLE: Primary structure of the Aequorea victorea green-f
; JOURNAL: Gene
; VOLUME: 111
; PAGES: 229-233
; DATE: 1992-01-01
; DATABASE ACCESSION NUMBER: Genbank M62653
; DATABASE ENTRY DATE: 1993-04-26
US-09-418-785-1

Query Match      98.9%;    Score 1262;   DB 4;   Length 238;
Best Local Similarity   98.7%;   Pred. No. 7.8e-130;
Matches  235;  Conservative  1;  Mismatches  1;  Indels  2;  Gaps  0;

Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

```
RESULT 13
US-09-129-192C-2
; Sequence 2, Application US/09129192C
; Patent No. 6495664
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AURO1270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/09/129,192C
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea
US-09-129-192C-2

Query Match      98.9%;    Score 1262;   DB 4;   Length 238;
Best Local Similarity   98.7%;   Pred. No. 7.8e-130;
Matches  235;  Conservative  1;  Mismatches  1;  Indels  2;  Gaps  0;

Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

```
RESULT 14
US-09-129-192C-74
; Sequence 74, Application US/09129192C
; Patent No. 6495664
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AURO1270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/09/129,192C
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 74
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea green fluorescent protein phosphorylation mutant
US-09-129-192C-74

Query Match      98.9%;    Score 1262;   DB 4;   Length 238;
Best Local Similarity   98.7%;   Pred. No. 7.8e-130;
Matches  235;  Conservative  1;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

```
RESULT 15
US-09-602-641-2
; Sequence 2, Application US/09602641
; Patent No. 6608189
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/602,641
; CURRENT FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/172,063
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-602-641-2

Query Match      98.9%;    Score 1262;   DB 4;   Length 238;
Best Local Similarity   98.7%;   Pred. No. 7.8e-130;
Matches  235;  Conservative  1;  Mismatches  2;  Indels  0;  Gaps  0;
```

```
Qy    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
        |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

Search completed: September 16, 2005, 17:22:10
Job time : 31 secs

THIS PAGE BLANK (USPTO)

```
            GenCore version 5.1.6

     Copyright (c) 1993 - 2005 Compugen Ltd.


OM protein - protein search, using sw model

Run on:     September 16, 2005, 17:20:18 ; Search time 113 Seconds
                                (without alignments)
                             852.856 Million cell updates/sec

Title:          US-10-757-624-4
Perfect score:  1276
Sequence:       1 MSKGEELFTGVVPILVELDG..........VLLEFVTAAGITHGMDELYK 238

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters:    1812044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :      Published Applications AA:*
        1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
        2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
        3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
        4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
        5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
        6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
        7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
        8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
        9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
       10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
       11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
       12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
       13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
       14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
       15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
       16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
       17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
       18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
       19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
       20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
       21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
       22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Match Length	DB	DB ID	Description	
1	1276	100.0	238	10	US-09-967-301-4	Sequence 4, Appli
2	1276	100.0	238	16	US-10-757-624-4	Sequence 4, Appli
3	1265	99.1	238	10	US-09-967-301-3	Sequence 3, Appli
4	1265	99.1	238	16	US-10-757-624-3	Sequence 3, Appli
5	1262	98.9	238	9	US-09-884-681-2	Sequence 2, Appli
6	1262	98.9	238	10	US-09-967-301-2	Sequence 2, Appli
7	1262	98.9	238	13	US-10-024-686-2	Sequence 2, Appli
8	1262	98.9	238	13	US-10-057-505-2	Sequence 2, Appli
9	1262	98.9	238	14	US-10-293-580-2	Sequence 2, Appli
10	1262	98.9	238	14	US-10-293-580-74	Sequence 74, Appl
11	1262	98.9	238	15	US-10-457-982-2	Sequence 2, Appli
12	1262	98.9	238	15	US-10-668-168-2	Sequence 2, Appli
13	1262	98.9	238	16	US-10-757-624-2	Sequence 2, Appli
14	1262	98.9	238	16	US-10-757-624-2	Sequence 2, Appli
15	1262	98.9	238	17	US-10-505-486-3	Sequence 3, Appli
16	1262	98.9	432	10	US-09-863-901-1	Sequence 1, Appli
17	1262	98.9	432	14	US-10-307-389-1	Sequence 1, Appli
18	1262	98.9	441	10	US-09-863-901-2	Sequence 2, Appli
19	1262	98.9	441	14	US-10-307-389-2	Sequence 2, Appli
20	1262	98.9	450	10	US-09-863-901-3	Sequence 3, Appli
21	1262	98.9	450	14	US-10-307-389-3	Sequence 3, Appli
22	1262	98.9	468	10	US-09-863-901-4	Sequence 4, Appli
23	1262	98.9	468	14	US-10-307-389-4	Sequence 5, Appli
24	1262	98.9	477	10	US-09-863-901-5	Sequence 5, Appli
25	1262	98.9	477	14	US-10-307-389-5	Sequence 6, Appli
26	1262	98.9	906	10	US-09-863-901-6	Sequence 6, Appli
27	1262	98.9	906	14	US-10-307-389-6	Sequence 34, Appl
28	1262	98.9	1070	10	US-10-001-486B-2	Sequence 4, Appli
29	1262	98.9	1099	14	US-10-259-864-4	Sequence 1, Appli
30	1262	98.9	1125	16	US-10-845-936A-34	Sequence 2, Appli
31	1262	98.9	1147	14	US-10-259-864-1	Sequence 2, Appli
32	1262	98.9	1452	13	US-10-050-673-2	Sequence 2, Appli
33	1259	98.7	238	9	US-09-899-954B-2	Sequence 6, Appli
34	1259	98.7	238	15	US-10-422-628-2	Sequence 4, Appli
35	1259	98.7	238	15	US-10-668-168-6	Sequence 1, Appli
36	1258	98.6	238	9	US-09-920-922-4	Sequence 125, App
37	1258	98.6	238	9	US-09-852-000-1	Sequence 2, Appli
38	1258	98.6	238	10	US-09-900-345A-125	Sequence 2, Appli
39	1258	98.6	238	10	US-09-866-538-2	Sequence 4, Appli
40	1258	98.6	238	10	US-09-794-308-2	Sequence 1, Appli
41	1258	98.6	238	10	US-09-865-291-2	Sequence 10, Appl
42	1258	98.6	238	14	US-10-121-258-10	Sequence 6, Appli
43	1258	98.6	238	14	US-10-221-461-6	Sequence 15, Appl
44	1258	98.6	238	14	US-10-321-195-15	Sequence 10, Appl
45	1258	98.6	238	14	US-10-305-765-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-967-301-4
; Sequence 4, Application US/09967301
; Publication No. US2003017585A1
; GENERAL INFORMATION:
;   APPLICANT: Stubbs, Simon L.
;   APPLICANT: Jones, Anne B.
;   APPLICANT: Michael, Nigel P.
;   APPLICANT: Thomas, Nicholas
;   TITLE OF INVENTION: Fluorescent Proteins
;   FILE REFERENCE: PA0111
;   CURRENT APPLICATION NUMBER: US/09/967,301
;   CURRENT FILING DATE: 2001-09-28
;   PRIOR APPLICATION NUMBER: GB 0109858.1
;   PRIOR FILING DATE: 2001-04-23
;   NUMBER OF SEQ ID NOS: 19
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: protein
US-09-967-301-4

  Query Match   100.0%; Score 1276; DB 10; Length 238;
  Best Local Similarity   100.0%; Pred. No. 5.6e-119;
  Matches 238; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

```
Qy    61  VTTLTYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  VTTLTYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 2
US-10-757-624-4
; Sequence 4, Application US/10757624
; Publication No. US20040138420A1
; GENERAL INFORMATION:
;   APPLICANT: Stubbs, Simon L. J.
;   APPLICANT: Jones, Anne E.
;   APPLICANT: Michael, Nigel P.
;   APPLICANT: Thomas, Nicholas
;   TITLE OF INVENTION: Fluorescent Proteins
;   FILE REFERENCE: PA0111
;   CURRENT APPLICATION NUMBER: US/10/757,624
;   CURRENT FILING DATE: 2004-01-14
;   PRIOR APPLICATION NUMBER: US 09/967,301
;   PRIOR FILING DATE: 2001-09-28
;   PRIOR APPLICATION NUMBER: GB 0109858.1
;   PRIOR FILING DATE: 2001-04-23
;   NUMBER OF SEQ ID NOS: 19
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic protein
US-10-757-624-4

Query Match             100.0%;  Score 1276; DB 16; Length 238;
Best Local Similarity   100.0%;  Pred. No. 5.6e-119;
Matches  238;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy    61  VTTLTYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  VTTLTYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 3
US-09-967-301-3
; Sequence 3, Application US/09967301
; Publication No. US20030175859A1
; GENERAL INFORMATION:
;   APPLICANT: Stubbs, Simon L.
;   APPLICANT: Jones, Anne E.
;   APPLICANT: Michael, Nigel P.
;   APPLICANT: Thomas, Nicholas
;   TITLE OF INVENTION: Fluorescent Proteins
;   FILE REFERENCE: PA0111
;   CURRENT APPLICATION NUMBER: US/09/967,301
;   CURRENT FILING DATE: 2001-09-28
;   PRIOR APPLICATION NUMBER: GB 0109858.1
;   PRIOR FILING DATE: 2001-04-23
;   NUMBER OF SEQ ID NOS: 19
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: protein
US-09-967-301-3

Query Match             99.1%;  Score 1265; DB 10; Length 238;
Best Local Similarity   99.2%;  Pred. No. 7e-118;
Matches  236;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy    61  VTTLTYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  VTTLSYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238


RESULT 4
US-10-757-624-3
; Sequence 3, Application US/10757624
; Publication No. US20040138420A1
; GENERAL INFORMATION:
;   APPLICANT: Stubbs, Simon L. J.
;   APPLICANT: Jones, Anne E.
;   APPLICANT: Michael, Nigel P.
;   APPLICANT: Thomas, Nicholas
;   TITLE OF INVENTION: Fluorescent Proteins
;   FILE REFERENCE: PA0111
;   CURRENT APPLICATION NUMBER: US/10/757,624
;   CURRENT FILING DATE: 2004-01-14
;   PRIOR APPLICATION NUMBER: US 09/967,301
;   PRIOR FILING DATE: 2001-09-28
;   PRIOR APPLICATION NUMBER: GB 0109858.1
;   PRIOR FILING DATE: 2001-04-23
;   NUMBER OF SEQ ID NOS: 19
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic protein
US-10-757-624-3

Query Match             99.1%;  Score 1265; DB 16; Length 238;
Best Local Similarity   99.2%;  Pred. No. 7e-118;
Matches  236;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy    61  VTTLTYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  VTTLSYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
```

Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 6
US-09-967-301-2
; Sequence 2, Application US/09967301
; Publication No. US20030175859A1
; GENERAL INFORMATION:
; APPLICANT: Stubbs, Simon L.
; APPLICANT: Jones, Anne E.
; APPLICANT: Michael, Nigel P.
; APPLICANT: Thomas, Nicholas
; TITLE OF INVENTION: Fluorescent Proteins
; FILE REFERENCE: PA0111
; CURRENT APPLICATION NUMBER: US/09/967,301
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: GB 0109858.1
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-967-301-2

Query Match 98.9%; Score 1262; DB 10; Length 238;
Best Local Similarity 98.7%; Pred. No. 1.4e-117;
Matches 235; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy 61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 7
US-10-024-686-2
; Sequence 2, Application US/10024686
; Publication No. US20020123113A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; Heim, Roger
; TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windiws95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/024,686
; FILING DATE: 17-Dec-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/057,995

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
Db 121 NRIELKGIDFKEDGNIIGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
Db 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITHGMDELYK 238

RESULT 5
US-09-884-681-2
; Sequence 2, Application US/09884681
; Patent No. US20020061546A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; Cubitt, Andrew B.
; TITLE OF INVENTION: Assays for Protein Kinases Using
; Fluorescent Protein Substrates
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/884,681
; FILING DATE: 19-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/679,865
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John S.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 023072-069000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-884-681-2

Query Match 98.9%; Score 1262; DB 9; Length 238;
Best Local Similarity 98.7%; Pred. No. 1.4e-117;
Matches 235; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy 61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGSVQLAD 180

Qy 181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

```
;     FILING DATE: <Unknown>
;     APPLICATION NUMBER: 08/727,452
;     FILING DATE: 10-OCT-1996
;     APPLICATION NUMBER: US95/14692
;     FILING DATE: 13-NOV-1995
;     APPLICATION NUMBER: 08/337,915
;     FILING DATE: 10-NOV-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Haile, Lisa A.
;     REGISTRATION NUMBER: 38,347
;     REFERENCE/DOCKET NUMBER: 07257/032002
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 619/678-5070
;     TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 238 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
;     FRAGMENT TYPE: internal
;     SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-024-686-2

Query Match             98.9%;  Score 1262;  DB 13;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 1.4e-117;
Matches   235;  Conservative   1;  Mismatches   2;  Indels   0;  Gaps   0;

Qy      1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60

Qy     61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           |||. :||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy    121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy    181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 8
US-10-057-505-2
; Sequence 2, Application US/10057505
; Publication No. US2002016467A1
; GENERAL INFORMATION:
;   APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
;   APPLICANT: AURORA BIOSCIENCES CORPORATION
;   APPLICANT: TSIEN, Roger
;   APPLICANT: HEIM, Roger
;   APPLICANT: CUBITT, Andrew
;   TITLE OF INVENTION: TANDEM FLUORESCENT PROTEIN CONSTRUCTS
;   FILE REFERENCE: REGEN1260-3
;   CURRENT APPLICATION NUMBER: US/10/057,505
;   CURRENT FILING DATE: 2002-01-25
;   PRIOR APPLICATION NUMBER: US 08/792,553
;   PRIOR FILING DATE: 1997-01-31
;   PRIOR APPLICATION NUMBER: US 09/396,003
;   PRIOR FILING DATE: 1999-09-13
;   NUMBER OF SEQ ID NOS: 31
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Aequorea victoria
US-10-057-505-2

Query Match             98.9%;  Score 1262;  DB 13;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 1.4e-117;
```

```
Matches   235;  Conservative   1;  Mismatches   2;  Indels   0;  Gaps   0;

Qy      1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60

Qy     61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           |||. :||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy    121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy    181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 9
US-10-293-580-2
; Sequence 2, Application US/10293580
; Publication No. US20030170767A1
; GENERAL INFORMATION:
;   APPLICANT: Aurora Biosciences Corporation
;   APPLICANT: Cubitt, Andrew B.
;   TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
;   FILE REFERENCE: AURO1270 (08366/031001)
;   CURRENT APPLICATION NUMBER: US/10/293,580
;   CURRENT FILING DATE: 2002-11-12
;   PRIOR APPLICATION NUMBER: US/09/129,192
;   PRIOR FILING DATE: 1998-07-24
;   NUMBER OF SEQ ID NOS: 74
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Aequorea
US-10-293-580-2

Query Match             98.9%;  Score 1262;  DB 14;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 1.4e-117;
Matches   235;  Conservative   1;  Mismatches   2;  Indels   0;  Gaps   0;

Qy      1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60

Qy     61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           |||. :||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy    121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGVQLAD 180

Qy    181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 10
US-10-293-580-74
; Sequence 74, Application US/10293580
; Publication No. US20030170767A1
; GENERAL INFORMATION:
;   APPLICANT: Aurora Biosciences Corporation
;   APPLICANT: Cubitt, Andrew B.
;   TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
;   FILE REFERENCE: AURO1270 (08366/031001)
;   CURRENT APPLICATION NUMBER: US/10/293,580
;   CURRENT FILING DATE: 2002-11-12
;   PRIOR APPLICATION NUMBER: US/09/129,192
```

; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 74
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea green fluorescent protein phosphorylation mutant
US-10-293-580-74

```
Query Match             98.9%;   Score 1262; DB 14;   Length 238;
Best Local Similarity   98.7%;   Pred. No. 1.4e-117;
Matches  235;  Conservative   1;  Mismatches   2;  Indels   0;  Gaps   0;

Qy     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL   60

Qy    61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV  120

Qy   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||
Db   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD  180

Qy   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK  238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK  238
```

RESULT 11
US-10-457-982-2
; Sequence 2, Application US/10457982
; Publication No. US2003021226SA1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/10/457,982
; CURRENT FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: US/09/602,641
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/172,063
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-10-457-982-2

```
Query Match             98.9%;   Score 1262; DB 15;   Length 238;
Best Local Similarity   98.7%;   Pred. No. 1.4e-117;
Matches  235;  Conservative   1;  Mismatches   2;  Indels   0;  Gaps   0;

Qy     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL   60

Qy    61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV  120

Qy   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||
Db   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD  180
```

```
Qy   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK  238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK  238
```

RESULT 12
US-10-668-168-2
; Sequence 2, Application US/10668168
; Publication No. US20040086968A1
; GENERAL INFORMATION:
; APPLICANT: Evans, Krista
; TITLE OF INVENTION: Mutants of Green Fluorescent Protein
; FILE REFERENCE: 0942.4020002
; CURRENT APPLICATION NUMBER: US/10/668,168
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/472,065
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/970,762
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: US 60/030,935
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria, gfp(h) S65T mutant cDNA clone
US-10-668-168-2

```
Query Match             98.9%;   Score 1262; DB 15;   Length 238;
Best Local Similarity   98.7%;   Pred. No. 1.4e-117;
Matches  235;  Conservative   1;  Mismatches   2;  Indels   0;  Gaps   0;

Qy     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL   60

Qy    61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV  120
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    61  VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV  120

Qy   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||
Db   121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD  180

Qy   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK  238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK  238
```

RESULT 13
US-10-724-178-2
; Sequence 2, Application US/10724178
; Publication No. US2004013752BA1
; GENERAL INFORMATION:
; APPLICANT: Odyssey Thera, Inc.
; APPLICANT: Michnick, Stephen
; APPLICANT: MacDonald, Marnie
; APPLICANT: Lamerdin, Jane
; TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT
; TITLE OF INVENTION: COMPLEMENTATION ASSAYS
; FILE REFERENCE: ODDY007
; CURRENT APPLICATION NUMBER: US/10/724,178
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US 60/461,133
; PRIOR FILING DATE: 2003-04-09
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-10-724-178-2

```
Query Match          98.9%;    Score 1262;    DB 16;    Length 238;
Best Local Similarity  98.7%;  Pred. No. 1.4e-117;
Matches  235;  Conservative  1;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60

Qy   61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||  :||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 14
US-10-757-624-2
; Sequence 2, Application US/10757624
; Publication No. US20040138420A1
; GENERAL INFORMATION:
; APPLICANT: Stubbs, Simon L. J.
; APPLICANT: Jones, Anne E.
; APPLICANT: Michael, Nigel P.
; APPLICANT: Thomas, Nicholas
; TITLE OF INVENTION: Fluorescent Proteins
; FILE REFERENCE: PA0111
; CURRENT APPLICATION NUMBER: US/10/757,624
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 09/967,301
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: GB 0109858.1
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-10-757-624-2

```
Query Match          98.9%;    Score 1262;    DB 16;    Length 238;
Best Local Similarity  98.7%;  Pred. No. 1.4e-117;
Matches  235;  Conservative  1;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60

Qy   61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||  :||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

RESULT 15
US-10-505-486-3
; Sequence 3, Application US/10505486
; Publication No. US20050118639A1
; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Determination of a ligand
; FILE REFERENCE: P03-0006PCT
; CURRENT APPLICATION NUMBER: US/10/505,486
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: JP 2002-45728
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: JP 2002-213949
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: JP 2002-298237
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 233
; SEQ ID NO 3
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea coerulescens
US-10-505-486-3

```
Query Match          98.9%;    Score 1262;    DB 17;    Length 238;
Best Local Similarity  98.7%;  Pred. No. 1.4e-117;
Matches  235;  Conservative  1;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL  60

Qy   61  VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         |||  :||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

Search completed: September 16, 2005, 17:39:32
Job time : 114 secs

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:       September 16, 2005, 17:06:47 ; Search time 24.5 Seconds
                                            (without alignments)
                                            934.677 Million cell updates/sec

Title:        US-10-757-624-4
Perfect score: 1276
Sequence:     1 MSKGEELFTGVVPILVELDG..........VLLEFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:        283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     PIR_79:*
               1: pir1:*
               2: pir2:*
               3: pir3:*
               4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1252	98.1	238	1	JQ1514	green-fluorescent
2	100	7.8	785	2	H72228	hypothetical prote
3	98.5	7.7	887	2	E82590	leucyl-tRNA synthe
4	92.5	7.2	534	1	NICLMA	nitrogenase (EC 1.
5	91	7.2	336	2	C64468	hypothetical prote
6	91	7.1	439	2	JH0414	synaptogamin o-p65
7	89	7.0	1224	1	ERHUAH	coatomer complex a
8	88.5	6.9	861	2	H64102	leucine-tRNA ligas
9	88	6.9	531	2	C95338	hypothetical prote
10	87.5	6.9	370	2	E70390	iron-sulfur cofact
11	87	6.8	357	2	G81155	tRNA (uracil-5-)-m
12	86.5	6.8	2222	1	A36028	DNA-directed DNA p
13	86.5	6.8	2573	2	D71614	hypothetical prote
14	86	6.7	281	2	AD2052	hypothetical prote
15	85.5	6.7	263	2	S53488	water-stress-induc
16	85.5	6.7	700	1	CIRUH2	calpain (EC 3.4.22
17	85.5	6.7	797	2	JC4078	protective surface
18	85.5	6.7	808	2	F64102	protective surface
19	85.5	6.7	860	2	AC0582	leucyl-tRNA synthe
20	85	6.7	471	2	T27856	hypothetical prote
21	85	6.7	632	2	T06586	DNA-binding protei
22	84.5	6.6	564	2	E81317	ABC-type transport
23	84.5	6.6	655	2	D83917	DNA topoisomerase
24	84.5	6.6	889	2	JC5576	inter-alpha-trypsi
25	83.5	6.5	422	2	B24815	calpain (EC 3.4.22
26	82.5	6.5	312	2	C81710	thioredoxin reduct
27	82.5	6.5	1134	1	A60234	IgA Fc receptor pr
28	82.5	6.5	1164	1	FCSOAG	IgA Fc receptor pr
29	82	6.4	470	2	T31049	hypothetical prote
30	81.5	6.4	390	2	C90208	hypothetical prote
31	81.5	6.4	613	2	A99552	oligoendopeptidase
32	81.5	6.4	636	2	C70031	hypothetical prote
33	81.5	6.4	788	1	JDVLHH	DNA-directed DNA p
34	81.5	6.4	2518	2	AI2140	polyketide synthas
35	81	6.3	422	1	BMRT2Y	synaptotagmin II -
36	81	6.3	647	2	H89988	hypothetical prote
37	81	6.3	874	2	JC4930	S-layer protein pr
38	80.5	6.3	389	2	G84245	NADH dehydrogenase
39	80.5	6.3	393	2	C64613	conserved hypothet
40	80.5	6.3	752	1	KXRTC1	proprotein convert
41	80.5	6.3	1259	2	AB1055	probable exported
42	80	6.3	353	2	E84941	imidazoleglycerol-
43	80	6.3	578	1	I40794	dihydrolipoamide d
44	80	6.3	760	2	E84715	probable oxysterol
45	80	6.3	764	2	T25012	hypothetical prote

ALIGNMENTS

RESULT 1
JQ1514
green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)
C;Species: Aequorea victoria
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
R;Prasher, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.
Gene 111, 229-233, 1992
A;Title: Primary structure of the Aequorea victoria green-fluorescent protein.
A;Reference number: JQ1514; MUID:92175527; PMID:1347277
A;Accession: JS0692
A;Molecule type: DNA
A;Residues: 1-107,'S',109-238 <PRA1>
A;Cross-references: UNIPROT:P42212; UNIPROT:Q17106; UNIPROT:Q17105; GB:M62654; NID:g1556
A;Accession: JQ1514
A;Molecule type: mRNA
A;Residues: 1-99,'F',101-140,'L',142-218,'V',220-238 <PRA2>
A;Cross-references: GB:M62653; NID:g155660; PIDN:AAA27721.1; PID:g155661
A;Accession: PQ0035
A;Molecule type: protein
A;Residues: 46-64;74-122;132-151;154-183;185-200 <PRA3>
R;Inouye, S.; Tsuji, F.I.
FEBS Lett. 351, 211-214, 1994
A;Title: Evidence for redox forms of the Aequorea green fluorescent protein.
A;Reference number: S48693; MUID:94364470; PMID:8082767
A;Accession: S48693
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-24,'Q',26-156,'P',158-171,'K',173-238 <INO>
A;Cross-references: GB:L29345; NID:g606383; PIDN:AAA58246.1; PID:g606384
R;Watkins, J.N.; Campbell, A.K.
submitted to the EMBL Data Library, January 1995
A;Reference number: S51330
A;Accession: S51330
A;Molecule type: mRNA
A;Residues: 1-13,'V',15-24,'Q',26-44,'N',46-153,'G',155-156,'P',158-171,'K',173-227,'R',
A;Cross-references: EMBL:X83959; NID:g634008; PIDN:CAA58789.1; PID:g634009
A;Experimental source: clone gfp1
A;Accession: S51331
A;Molecule type: mRNA
A;Residues: 1-24,'Q',26-29,'R',31-83,'L',85-153,'G',155-156,'P',158-171,'K',173-208,'Q',
A;Cross-references: EMBL:X83960; NID:g634010; PIDN:CAA58790.1; PID:g634011
A;Experimental source: clone gfp2
R;Yang, F.; Moss, L.G.; Phillips Jr., G.N.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65692; PDB:1GFL
A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 'A',2-79,'R',81-9
A;Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli
R;Yang, F.; Moss, L.G.; Phillips Jr., G.N.
Nat. Biotechnol. 14, 1246-1251, 1996
A;Title: The molecular structure of green fluorescent protein.
A;Reference number: A58953; MUID:98294543; PMID:9631087

```
A;Contents: annotation; X-ray crystallography, 1.9 angstroms
C;Comment: This protein is excited by the photoprotein aequorin (see PIR:AQJFNV) emittin
C;Genetics:
C;Gene: GFP
A;Introns: 69/3; 167/3
C;Superfamily: green-fluorescent protein
C;Keywords: chromoprotein; luminescence
F;65-67/Cross-link: 5-imidazolinone (Ser-Gly) #status experimental
F;66/Modified site: dehydrotyrosine (Tyr) #status experimental

Query Match     98.1%;     Score 1252; DB 1;  Length 238;
Best Local Similarity     97.1%;   Pred. No. 3.1e-99;
Matches  231;  Conservative    5;   Mismatches    2;   Indels    0;  Gaps    0;

Qy    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61  VTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
         ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121  NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
         ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  121  NRIELKGIDFKEDGNILGHKMEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
         |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db  181  HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMILLEFVTAAGITHGMDELYK 238


RESULT 2
H72228
hypothetical protein TM1624 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: H72228
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-785 <ARN>
A;Cross-references: UNIPROT:Q9XIV9; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AAD3669
A;Experimental source: strain MSB8
C;Genetics:
C;Gene: TM1624

Query Match     7.8%;     Score 100; DB 2;  Length 785;
Best Local Similarity     20.9%;   Pred. No. 1.8;
Matches   40;  Conservative   27;   Mismatches   27;   Indels   64;  Gaps    5;

Qy    2  SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 61
         : |  |||| :|:|   |||
Db   15  NEGRFSFEGTVPGVVQAD----------------------------LVRKGLLPHPYVGM- 46

Qy   62  TTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVN 121
         ::||||                          ::
Db   47  --------------NEDLFKEIEDREWIYEREFEFKEDVKEGERVDLVFEGVDTLS 88

Qy  122  RIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLADH 181
         :|:|: |||   |  :|| |||  :|||
Db   89  DVYLNGVYL--GSTBDMFIEYRFDVTNVL---KEKNHLKVYIK------SPIRVPKT 134

Qy  182  YQQNTPIGDGP 192
         ||: : |||
Db  135  LEQNYGVLGGP 145
```

```
RESULT 3
E82590
leucyl-tRNA synthetase XF2176 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: E82590
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:1091034 7
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: E82590
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-887 <SIM>
A;Cross-references: GB:AE004031; GB:AE003849; NID:g9107309; PIDN:AAF84975.1; GSPDB:GN001:
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.V.; Santelli, R.V.; Sawasaki
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Za
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
C;Gene: XF2176
C;Superfamily: leucine-tRNA ligase

Query Match     7.7%;     Score 98.5; DB 2;  Length 887;
Best Local Similarity     23.2%;   Pred. No. 2.8;
Matches   46;  Conservative   27;   Mismatches   68;   Indels   57;  Gaps   10;

Qy   49  TTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGVVQERTIFFKDDGNY-- 106
         :||  :  |
Db  329  TNEQLPV-WVANFVLMAYGTGAVMAVPGHDQDRDQEF--ANKYGLPIRQVIALKEPKNQDE 385

Qy  107  ------------------------------KTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYI 152
                                       |||  |
Db  386  STWEPDVWRDWYADKTR---EFE---LINSAEFDGLDYQDAFEVLAERFE--------- 429

Qy  153  MADKQKNG-IKVNFKIRHNIEDGGVQLADHYQQNTPI------------- 198
         |  :: |   :
Db  430  ---RQGRGQRRVNYRLR----DWGVSRQRYWGCPIPVIYCPTCGAVPVPEDQLPVLPEN 482

Qy  199  -HYLSTQSALSKDPNEKR 215
         :  : ::
Db  483  VAFSGTGSPIKTDPEWRK 500


RESULT 4
NICLMA
nitrogenase (EC 1.18.6.1) molybdenum-iron protein alpha chain [validated] - Clostridium p
N;Alternate names: dinitrogenase alpha chain; nitrogenase component I alpha chain
C;Species: Clostridium pasteurianum
C;Date: 01-Sep-1981 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C;Accession: S07389; I40816; A00545
R;Wang, S.Z.; Chen, J.S.; Johnson, J.L.
Nucleic Acids Res. 15, 3935, 1987
A;Title: Nucleotide and deduced amino acid sequences of nifD encoding the alpha-subunit
A;Reference number: S07389; MUID:87231095; PMID:3473447
A;Accession: S07389
A;Molecule type: DNA
A;Residues: 1-534 <WAN>
A;Cross-references: UNIPROT:P00467; EMBL:Y00155; NID:g40583; PIDN:CAA68349.1; PID:g58099:
R;Chen, K.C.
```

J. Bacteriol. 166, 162-172, 1986
A;Title: Structural features of multiple nifH-like sequences and very biased codon usage
A;Reference number: I40814; MUID:8616010; PMID:3457003
A;Accession: I40816
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-167 <CHB>
A;Cross-references: GB:M21537; NID:g144870; PIDN:AAA83531.1; PID:g551775
R;Hase, T.; Nakano, T.; Matsubara, H.; Zumft, W.G.
J. Biochem. 90, 295-298, 1981
A;Title: Correspondence of the larger subunit of the MoFe-protein in clostridial nitroge
A;Reference number: A00545; MUID:82030699; PMID:7026551
A;Accession: A00545
A;Molecule type: protein
A;Residues: 2-41,'K',43-94,'D',96-180 <HAS>
R;Kim, J.; Woo, D.; Rees, D.C.
submitted to the Brookhaven Protein Data Bank, March 1993
A;Reference number: AS1301; PDB:1MIO
A;Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 2-425,'A',427-526
C;Comment: The key enzymatic reactions in nitrogen fixation are catalyzed by the nitroge
and the iron protein (also called component II or nitrogenase reductase).
C;Comment: Each alpha/beta dimer covalently binds one molybdenum-iron-sulfur cluster and
C;Genetics:
A;Gene: nifD
A;Start codon: GTG
C;Complex: heterotetramer of two alpha and two beta chains
C;Function:
A;Description: the enzyme complex catalyzes the reduction of dinitrogen to 2 molecules o
A;Pathway: nitrogen fixation
C;Superfamily: dinitrogenase alpha chain; nitrogenase vanadium-iron protein alpha chain
C;Keywords: 4Fe-4S; ATP; heterotetramer; iron-sulfur protein; metalloprotein; molybdenum
F;2-534/Product: nitrogenase molybdenum-iron protein alpha chain experimental <M
F;5-522/Domain: nitrogenase vanadium-iron protein alpha chain homology <VIA>
F;53,79,145/Binding site: 4Fe-4S cluster 1 (Cys) (covalent) #status experimental
F;79/Binding site: 4Fe-4S cluster 2 (Cys) (covalent) #status experimental
F;262/Binding site: homocitryl Mo-7Fe-8S cluster (Cys) experimental
F;482/Binding site: homocitryl Mo-7Fe-8S cluster molybdenum (His) (ligand) #status exper

Query Match 7.2%; Score 92.5; DB 1; Length 534;
Best Local Similarity 19.9%; Pred. No. 4.7;
Matches 48; Conservative 39; Mismatches 97; Indels 57; Gaps 8;

```
Qy  17 ELDGDVNGHKFSVSGEGEGDATYGKL------------TLKFIC--TTGKLPVPW-- 57
        |:|:|||||                              |||||  ||||||||
Db 224 EMDRVLEKIGYHVNATLTGDATYEKVQNADKADLNLVQCHRSINYIABMMETKYGIPWIK 283

Qy  58 -----PTLVTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNY-KTRA 110
             ||||:||     ||  |:|:||||
Db 284 CNFIGVDGIVETLRDMAKCFDD-PELTKR------------TEBVIAEBIAAIQDDLDYFKBKL 334

Qy 111 EVK--------FEGDTLVNRIELKGIDFKEDGNILGHKLEY---------------NYNSH 148
        |||        :|:|| ::|
Db 335 QGKTACLYVGGSRSHTYMNMLKSFGVDSLVAGFEFAHRDDYEGREBVIPTIKIDADSKNIP 394

Qy 149 NVYIMADKQKNGIKVNFKIRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS 208
         ||:|     ||:|||
Db 395 EITVTPEDQKYRVVIPEDKVEBLKKAGVPLSSYGGMMKEMHDGTILIDDMNHHDMEVVLE 454

Qy 209 K 209
        |
Db 455 K 455
```

RESULT 5
C64468
hypothetical protein MJ1348 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: C64468
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.
.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996

A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: C64468
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-336 <BUL>
A;Cross-references: UNIPROT:Q58743; GB:U67574; GB:L77117; NID:g1591978; PIDN:AAB99360.1;
C;Genetics:
A;Map position: FOR1295121-1296131
A;Start codon: TTG

Query Match 7.2%; Score 92; DB 2; Length 336;
Best Local Similarity 24.9%; Pred. No. 2.8;
Matches 66; Conservative 36; Mismatches 81; Indels 82; Gaps 17;

```
Qy  19 DGDVNGHKFSVSGE-------GEGDATYGKL--TLKFICTGKLPVPW----------PTL 60
        |:|                  |||||:|||    |||:| |||||||
Db  91 DGDY----YNLSGELSTIASIFAKIGKLDITKNFLNSGGE----WIYNDVIKDANSEDIL 142

Qy  61 VTTLTYGV--QCFSRYPDHMKR------------HDFFKSAMPEGYVQERTIFFKDDG 104
                  |                 :|||
Db 143 KSVLTFDFSYEBRKEILNREP-HLRKLPEDNIYNFNSDPFDFMM---------MFFIGAG 192

Qy 105 NYKTRAEVKFE------------GDTLVNRIELKGIDFKEDGNIL--GHKLEYNYNSHNVYI 152
        |:|                    :||||:||||||||:
Db 193 NWKRFLEVYEEFKNKIKSCQISNEIVNEI-IKRFD-KSMSDLLAIAHLLKENYEKCLYYV 250

Qy 153 MADK----------------QKNGIKVNFKIRHNIEDGGVQLADHYQQNTPIGDGPVLLP-DNFY 200
        || |                ||||:|||||||||||
Db 251 MLFKEYFDLDBFNEIEKNKINLIVDIAYNLKNGNVKKEEWLNRLNEIYKEIIKRPLPNTY 310

Qy 201 LSTQSALSKD-PNEKRDHMVLLEFV 224
             | |   |:|:|||| ||:|
Db 311 ---KDAHNDLLNEILDYYVLKEFI 332
```

RESULT 6
JH0414
synaptogamin o-p65-B - electric ray (Discopyge ommata)
N;Alternate names: synaptic vesicle protein o-p65-B
C;Species: Discopyge ommata
C;Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 09-Jul-2004
C;Accession: JH0414; PS0223
R;Wendland, B.; Miller, K.G.; Schilling, J.; Scheller, R.H.
Neuron 6, 993-1007, 1991
A;Title: Differential expression of the p65 gene family.
A;Reference number: JH0413; MUID:91273991; PMID:2054189
A;Molecule type: mRNA
A;Residues: 1-439 <WEN>
A;Cross-references: UNIPROT:P24506; GB:M64276; NID:g213110; PIDN:AAA49228.1; PID:g213111
A;Experimental source: electric organ
A;Accession: PS0223
A;Molecule type: protein
A;Residues: 'MLV',26-34;'XX',194-199,'X',201-206;'X',322-332,'D',334-337 <WEN1>
C;Superfamily: synaptotagmin; protein kinase C C2 region homology
C;Keywords: glycoprotein; membrane protein; synaptic vesicle
F;75-101/Domain: hydrophobic <HYD>
F;153-266/Domain: protein kinase C C2 region homology <KC2A>
F;284-399/Domain: protein kinase C C2 region homology <KC2B>
F;6,46/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.1%; Score 91; DB 2; Length 439;
Best Local Similarity 20.1%; Pred. No. 4.9;
Matches 58; Conservative 44; Mismatches 92; Indels 94; Gaps 13;

```
Qy  15 LVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPV-PWPTLVTTLTYGV----- 68
          |:|||   |  ||:|||||
Db  38 MMPIDTGDNSTRAGVPGGEGKND-VFEKLKEKFWNBLQKIPLPPWALIAIAIVSGLLLTC 96

Qy  69 ------QCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG------- 116
```

RESULT 8
H64102
leucine-tRNA ligase (EC 6.1.1.4) - Haemophilus influenzae (strain Rd KW20)
N;Alternate names: leucyl-tRNA synthetase
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C;Accession: H64102
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: H64102
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-861 <TIGR>
A;Cross-references: UNIPROT:P43827; GB:U32774; NID:g1573942; PIDN:AAC22581.1;
C;Genetics:
A;Gene: leuS
C;Superfamily: leucine-tRNA ligase
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match　　　6.9%;　　Score 88.5;　　DB 2;　　Length 861;
Best Local Similarity　24.3%;　　Pred. No. 19;
Matches　45;　Conservative　26;　Mismatches　67;　Indels　47;　Gaps　9;

Qy　　49　TTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKT　108
　　　　　|| |||　|| :|||||||||
Db　　314　TGDKLPI-WVANFVLMHYGTGAVMAVPAHDQRD-FEFAQKYSLPIKQVIAPLADEEIDL　370

Qy　　109　RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADK-QKNGI--KVN　164
　　　　　|||　||　　:||:
Db　　371　TKQAFVEHGKLVNSDEFDGKNF--DGAFNG-------IADKLEKLGVGKRQVN　414

Qy　　165　FKIRH------NIEDGGVQLADHYQQNTPIGDGPVLLPDNHYL-STQSAL　207
　　　　　　:|||||:　:||:||　:||:||
Db　　415　YRLRDWGVSRQRYWGAPIPMLTLENGDVVPA------PMEDLPIILPEDVVMDGVKSPI　467

Qy　　208　SKDPN　212
　　　　　:|||
Db　　468　NADPN　472

RESULT 9
C95338
hypothetical protein SMa1131 [imported] - Sinorhizobium meliloti (strain 1021) megaplasmi
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95338
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowse
; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: C95338
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-531 <KUR>
A;Cross-references: UNIPROT:Q92Z81; GB:AB006469; PIDN:AAK65269.1; PID:g14523721; GSPDB:GN
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galiber, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.(
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:

Db　　97　CLCICKKCCKKKKKNKKEKGKGK-----------------KNDINMK---DVKGSGNQDDD　138

Qy　　117　--DTLVNRIELKGIDFKEDGNI--LGHKLEYNYNSH------------------　148
　　　　　:||:　　　||

Db　　139　DAETGLTEGEDKEBEAKEEEKLGKIQFSLDYDFQANQLITVGIIQAAELPALDMGGTSDPY　198

Qy　　149　-NVYIMADKQKN-GIKVN-------------FKIRHNIEDGGVQLA-------DHYQQNTP　187
　　　　　|||：||：||

Db　　199　VKVFLLPDKKKKYETKVQKTLNPFNESFVFKVPYQ-BLGGKTLMMAVYDFDRFSKHDC　257

Qy　　188　IGDGPVLLPD-------NHYLSTQSALSKDPNEKRDHMVLLEFVTAAG　228
　　　　　|||：　　　　　　||

Db　　258　IGQVTVLMTKVDLGQQLEEWRDLBSAEKEEPEKLGDICTSLRVVPTAG　305

RESULT 7
ERHUAH
coatomer complex alpha chain homolog - human
N;Alternate names: HEP-COP; xenopsin homolog; xenopsin-related peptide precursor
N;Contains: xenin 25; xenopsin-related peptide
C;Species: Homo sapiens (man)
C;Date: 10-May-1996 #sequence_revision 08-Nov-1996 #text_change 09-Jul-2004
C;Accession: JC4668; A44317
R;Chow, V.T.K.; Quek, H.H.
Gene 169, 223-227, 1996
A;Title: HEP-COP, a novel human gene whose product is highly homologous to the alpha-sub
A;Reference number: JC4668; MUID:96194806; PMID:8647451
A;Accession: JC4668
A;Molecule type: mRNA
A;Residues: 1-1224 <CHO>
A;Cross-references: UNIPROT:P53621; GB:U24105; NID:g1638873; PIDN:AAB70879.1; PID:g10023
A;Experimental source: Hep3B hepatocellular carcinoma cell
R;Feurle, G.E.; Hamscher, G.; Kusiek, R.; Meyer, H.E.; Metzger, J.W.
J. Biol. Chem. 267, 22305-22310, 1992
A;Title: Identification of xenin, a xenopsin-related peptide, in the human gastric mucos
A;Reference number: A44317; MUID:93054515; PMID:1429581
A;Accession: A44317
A;Molecule type: protein
A;Residues: 1-25 <FBU>
A;Experimental source: gastric mucosa
A;Note: sequence extracted from NCBI backbone (NCBIP:117018)
A;Note: plasma levels of xenin 25, as determined by immunoassay, rise after meals and it
C;Genetics:
A;Gene: GDB:COPA; HEP-COP
A;Cross-references: GDB:4642787; OMIM:601924
A;Map position: 1q23-1q25
C;Superfamily: yeast coatomer complex alpha chain; WD repeat homology
C;Keywords: duplication; hormone; plasma; stomach
F;1-25/Product: xenin 25 #status experimental <XNP>
F;5-38/Domain: WD repeat homology <WD1>
F;17-25/Product: xenopsin-related peptide #status predicted <XRP>
F;47-80/Domain: WD repeat homology <WD2>
F;89-122/Domain: WD repeat homology <WD3>
F;131-164/Domain: WD repeat homology <WD4>
F;201-234/Domain: WD repeat homology <WD5>
F;245-278/Domain: WD repeat homology <WD6>

Query Match　　　7.0%;　　Score 89;　　DB 1;　　Length 1224;
Best Local Similarity　25.9%;　　Pred. No. 28;
Matches　41;　Conservative　26;　Mismatches　59;　Indels　32;　Gaps　7;

Qy　　90　EGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIE-LKGIDFKEDGNILGHKLEYNYNSH　148
　　　　　|||:　|||：　||　||：　：||||||||||||
Db　　632　KGYPEVALHFVKDE--KTRFSLALECGNIEIALEAAKALDDKNCWEKLGEVALLQGNHQ　688

Qy　　149　NVYIMADKQKNGIKVNF---------KIRHNIEDGGVQ-------LADHYQONTPIGDGPVLLPD　197
　　　　　　::　::|：　||　：　：　|||||
Db　　689　IVEMCYQRTKNFDKVSFLYLITGNLEKLRKMMKIABIRKDMSGHYQNALYLGD-------741

Qy　　198　NHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDE　235
　　　　　　:|：：|||||:|||||
Db　　742　---VSERVRILKNCGQKS-----LAYLTAA--THGLDE　769

RESULT 11
G81355
tRNA (uracil-5-)-methyltransferase (EC 2.1.1.35) Cj0831c [imported] - Campylobacter jeju
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: G81355
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: G81355
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-357 <PAR>
A;Cross-references: UNIPROT:Q9PP92; GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB7309
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: trmA; Cj0831c
C;Keywords: methyltransferase; S-adenosylmethionine

Query Match   6.8%; Score 87; DB 2; Length 357;
Best Local Similarity 24.8%; Pred. No. 8.2;
Matches 30; Conservative 18; Mismatches 39; Indels 34; Gaps 5;

Qy   79 KRHDFFKSAMPEGVVQERTIFFKDDGNYKTRAEVKF-EGDTLV-------NRIELKG 127
               ::|:||||||||::|||||||:::||||:: :| .: : :::   .::
Db   14 EKHSFIKYFKEFYTKDFKLFASKDKHYRTRAELSFYHENDTLFYAMEDPKSKKYIIEY 73

Qy   128 IDFKBD--------GNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIE 172
               :| ::| ||||||||||||||||||:||||||
Db   74 LDFADEKICAFMPRLLEYLRQDNKLKEKL-----FGVEFLTKQE-LSITLLYHKNIE 125

Qy   173 D 173
           |
Db   126 D 126

RESULT 12
A36028
DNA-directed DNA polymerase (EC 2.7.7.7) II catalytic chain - yeast (Saccharomyces cerev
N;Alternate names: DNA-directed DNA polymerase II chain A; protein NO825; protein YNL262
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A36028; B36028; S60919; S63235; S65121
R;Morrison, A.; Araki, H.; Clark, A.B.; Hamatake, R.K.; Sugino, A.
Cell 62, 1143-1151, 1990
A;Title: A third essential DNA polymerase in Saccharomyces cerevisiae.
A;Reference number: A36028; MUID:90381771; PMID:2169349
A;Accession: A36028
A;Molecule type: protein
A;Residues: 1-2222 <MOR>
A;Cross-references: UNIPROT:P21951; GB:M60416; GB:M36724; NID:g171408; PIDN:AAA88711.1;
A;Accession: B36028
A;Molecule type: protein
A;Residues: 1214-1216,'X',1218-1221 <MO2>
R;Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of a 24152 bp segment from the left arm of chromosome XIV fr
A;Reference number: S60909
A;Accession: S60919
A;Molecule type: DNA
A;Residues: 1-2221 <SEN>
A;Cross-references: EMBL:X92494; NID:g1045236; NID:g1045247; PIDN:CAA63235.1; PID:g1045247
R;Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63235
A;Accession: S63235
A;Molecule type: DNA
A;Residues: 1-2222 <SEW>
A;Cross-references: EMBL:Z71538; NID:g1302316; NID:g1302317; PIDN:CAA96169.1; PID:g1302317; GSPDB:GN00
A;Experimental source: strain S288C

A;Gene: SMal131
A;Genome: plasmid

Query Match   6.9%; Score 88; DB 2; Length 531;
Best Local Similarity 20.4%; Pred. No. 11;
Matches 61; Conservative 29; Mismatches 87; Indels 122; Gaps 14;

Qy   6 ELFTGVPILVELDGDVN-GHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTL 64
           |:|:||| ||:||||||: ||
Db   176 EIGTGPSPLRLLFSGDIGPRHKL------------LQF--DAKAPSGWDYVICES 216

Qy   65 TYGVQCFSRYPDHMKRH----DFFKSAMPEG------YVQERT------IFFKD---- 102
           ||||||||||||||||| :|:|||||||| ::|||| |||||
Db   217 TYGNIEREEADDAARRHILRSEVLTAAHPNGALLIPSFAVERTQELLTDLVHLMETGAVP 276

Qy   103 -----DGNYKTRAEVKF--------EGDTLVNRIELKGIDFKB--------- 132
                |||||||||||| ||||||||||||:|||
Db   277 KCPIIIDSPLATRASEIFRRHARELENGDALVRGIESKNVRFTETAEQSKAVDLIRGFHI 336

Qy   133 -------DGNILGHKLE-YNYNSHNVYIMADKQKNG----------------IKVN 164
                  ||||||||: |||||||||||||||| |||:
Db   337 VIAASGMCEAGRIRHRLKNWLWRDEATVLLVGYQANGTLGRILLEBGATVVRIQGDDIRVR 396

Qy   165 FKIRH-NIEDG---GVQLADHYQQNTPIGDGPVL---------------LPDNHYL 201
           |:||| ||::   ::|||||:|||: :.:| ||||||:|
Db   397 ARIRKLDIYSGHADGSELADWVRARQPISAGVFLVHGEDEALEGIRERLSTFLPQDHLI 455

RESULT 10
E70390
iron-sulfur cofactor synthesis protein nifs - Aquifex aeolicus
C;Species: Aquifex aeolicus
N;Contains: L-cysteine sulfurtransferase (EC 2.8.1.-)
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: E70390
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
v.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98166666; PMID:9537320
A;Accession: E70390
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-370 <AQF>
A;Cross-references: UNIPROT:O67152; GB:AE000720; NID:g2983529; PIDN:AAC07111.1; PID:g298
A;Experimental source: strain VF5
C;Genetics:
A;Gene: nifS1
C;Superfamily: nitrogen fixation protein nifs
C;Keywords: phosphoprotein; pyridoxal phosphate; sulfurtransferase
C;Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
F;195/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
F;318/Active site: Cys (cysteine persulfide intermediate) #status predicted

Query Match   6.9%; Score 87.5; DB 2; Length 370;
Best Local Similarity 25.4%; Pred. No. 7.8;
Matches 49; Conservative 30; Mismatches 91; Indels 23; Gaps 8;

Qy   3 KGBELFTGVV----PILVELD---GDVNGHKF-SVSGEG----EGDATYGKLTLKFICT 49
           |||||||| |:|:| ||||||| ||| || |||||||||||||||
Db   164 KGVFLLTDAVQAIGKIPIELKNISYATFSGHKFHAIKGSGFLYISDEANYEPLIVGGGQE 223

Qy   50 TGKLP-----VPWPTLVTTLTYGVQCFSRYPDHMKR-HDFFKSAMPEGVVQERTIFFKDD 103
           |||| ||||||||||||||||||||||||| |||||||||||||||||||||||
Db   224 NGKRSGTENVGILSLAKALBIIVSNFSRYQEQLKKLRDLFENLLLEA-LPDAQIVGKDA 282

Qy   104 GNYKTRABV--KFEGDTLVNRIELKGIDFKKEDGNILGHKLEYNYNSHNVYIMADKQKNG 160
           ||||||| :|||||||||||||||| :||||||||||||||||||||||||||||||
Db   283 ERSPSISSVIMPKFFGAEIVNKLSEKGIYCSTGSACLSGEYEPNKHMLKWGFSQEKALRM 342

Qy   161 IKVNPKIRHNIED 173
           |||:||:|||:||
Db   343 VRFSFGLLNKEEE 355

R;Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.
Yeast 12, 505-514, 1996
A;Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from Sa
A;Reference number: S65111; MUID:96310631; PMID:8740425
A;Accession: S65121
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2221 <SEF>
A;Cross-references: EMBL:X92494; NID:g1045236; PIDN:CAA63235.1; PID:g1045247
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Gene: SGD:POL2; DUN2; MIPS:YNL262w
A;Cross-references: SGD:S0005206; MIPS:YNL262w
A;Map position: 14L
A;Superfamily: DNA-directed DNA polymerase II
C;Keywords: DNA binding; nucleotidyltransferase; nucleus; zinc finger

Query Match   28.2%; Score 86.5; DB 1; Length 2222;
Best Local Similarity
Matches 37; Conservative 14; Mismatches 49; Indels 31; Gaps 7;

Qy   53 LPVPWP-TLVTLTYGVQCFSRYPDHM--------KRHDFFKSAMPEGYV----QERTI   98
Db   883 LPKSFPETYFFTLENGKKLYLSYPCSMLNRVRHQKFTNHQYQELKDPLNVIYETHSENTI 942

Qy   99 FFKDDGNYKTR--AEVKFEGDTLVNR---------IELKGIDFKEDGNILGHKLEYNYN 146
Db   943 FFEVDGPYKAMILPSSKEEGKGIKKRYAVFNEDGSLAELKGFELKRRGEL---QLIKNFQ 999

Qy   147 S--HNVYIMAD 155
Db   1000 SDIFKVFLEGD 1010

RESULT 13
D71614
hypothetical protein PFB0460c - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C;Accession: D71614
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.
.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
.; Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A;Reference number: A71600; MUID:99021743; PMID:9804551
A;Accession: D71614
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2573 <GAR>
A;Cross-references: UNIPROT:O96185; GB:AE001396; GB:AE001362; NID:g3845188; PIDN:AAC7188
A;Experimental source: clone 3D7
C;Genetics:
A;Gene: PFB0460c

Query Match 26.2%; Score 86.5; DB 2; Length 2573;
Best Local Similarity
Matches 34; Conservative 29; Mismatches 54; Indels 13; Gaps 5;

Qy   93 VQERTIFFKD--DGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYN--YNSH 148
Db   126 LKKETLLCKDIKSGSSNDPMDEISLFKDDMVDDKELK--DFEKSSLKIKNKEVYNFIYNKM 183

Qy   149 NVYIMADKQKNGIKVNFKIRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS 208
Db   184 NLHIKENKKKDEKEKKNKIHHNDENNNM----IYYKNI---DKTHYILDNNVVHILNDIN 236

Qy   209 KDPNEKRDHM 218
Db   237 TYLKRERDYM 246

RESULT 14
AD2052

hypothetical protein all1970 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AD2052
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anab
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2052
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-281 <KUR>
A;Cross-references: UNIPROT:Q8YVK5; GB:BA000019; PIDN:BAB73669.1; PID:g17131060; GSPDB:GN
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all1970

Query Match 24.4%; Score 86; DB 2; Length 281;
Best Local Similarity
Matches 33; Conservative 26; Mismatches 54; Indels 22; Gaps 6;

Qy   106 YKTRAEVKFEGDTLVNRIELKGI---DFKEDGNILGHKLEYNYNSHNVYIMADKQKNGI 161
Db   46 FKARS-LQSSDILLNPIKYNLQNRPDWKKDEYYIVHAEGYKYKHE-----SSKRKHSI   99

Qy   162 KVNFKI--RHNIEDGGVQLADH--------YQQNTPIGDG--PVLLPDNHYLSTQSALSK 209
Db   100 KPGFSLQERGDIVNQVLQLLDNVLIPDEPIQTDIFTPSGQNTPLAMRDYDYKFISQPKTITK 159

Qy   210 DPNEKRDHMVLLEFV 224
Db   160 QQKEKLSQDKIQELV 174

RESULT 15
S53488
water-stress-inducible protein DS2 - Chaco potato
C;Species: Solanum chacoense (Chaco potato)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S53488
R;Silhavy, D.; Hutvagner, G.; Barta, E.; Banfalvi, Z.
Plant Mol. Biol. 27, 587-595, 1995
A;Title: Isolation and characterization of a water-stress-inducible cDNA clone from Solan
A;Reference number: S53488; MUID:95201251; PMID:7894021
A;Accession: S53488
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-263 <SIL>
A;Cross-references: UNIPROT:Q41300; GB:U12439; NID:g607904; PIDN:AAA86052.1; PID:g607905

Query Match 22.3%; Score 85.5; DB 2; Length 263;
Best Local Similarity
Matches 46; Conservative 21; Mismatches 88; Indels 51; Gaps 6;

Qy   31 GEGEGDATYGKLTLKPICTGKLPVWPTLVTTLTYGVQCFSRYPDHMKRHDFFKSAMPE   90
Db   31 GESEKTSYGEKTSYGDDTYGB------------KTTTFGDD--NKYGEKTSYGDDTYGEKPT   79

Qy   91 GY-----VQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFFKEDGNILGHKLEY-- 143
Db   80 SYGGDNTYGEKTSYGKGDDNKYGEKTSYGEG-------------DDNKYGEKTSYGD 123

Qy   144 -NYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLS 202
Db   124 SGYGEKPSYGGGDDNKYGEKTSYGNEEGGYGGGVGETTNYBEN----------------- 167

Qy   203 TQSALSKDPNEKRDHMVLLEFVTAAG 228
Db   168 SETKTSBDYKBEKKHHKHLEEIGGLG 193

us-10-757-624-4.rpr

Search completed: September 16, 2005, 17:21:03
Job time : 25.5 secs

THIS PAGE BLANK (USPTO)

```
               GenCore version 5.1.6
      Copyright (c) 1993 - 2005   Compugen Ltd.

OM protein - protein search, using sw model

Run on:         September 16, 2005, 16:55:06 ; Search time 115 seconds
                                  (without alignments)
                                  1059.781 Million cell updates/sec

Title:          US-10-757-624-4
Perfect score:  1276
Sequence:       1 MSKGEELFTGVVPILVELDG.........VLLEFVTAAGITHGMDELYK 238

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:      1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      UniProt 03:*
                1: uniprot_sprot:*
                2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                           SUMMARIES
```

Result No.	Score	Query Match	Length	DB	ID	Description
1	1262	98.9	238	2	Q8GHE2	Q8ghe2 azotobacter
2	1258	98.6	238	1	GFP_AEQVI	P42212 aequorea vi
3	1258	98.6	238	2	Q71RY9	Q71ry9 azotobacter
4	1257	98.5	238	2	Q8GHE4	Q8ghe4 azomonas ag
5	1256	98.4	238	2	Q8GHE3	Q8ghe3 azotobacter
6	1252	98.1	238	2	Q93125	Q93125 aequorea vi
7	1216	95.3	238	2	Q17105	Q17105 aequorea vi
8	1201	94.1	238	2	Q17106	Q17106 aequorea vi
9	1188	93.1	238	2	Q6YGZ0	Q6ygz0 aequorea co
10	1109	86.6	238	2	Q8WTC6	Q8wtc6 aequorea ma
11	1105	86.6	238	2	Q8WP95	Q8wp95 aequorea ma
12	1102	86.4	238	2	Q8WTC4	Q8wtc4 aequorea ma
13	1100	86.2	238	2	Q8WTD0	Q8wtd0 aequorea ma
14	1099	86.1	238	2	Q8WTC8	Q8wtc8 aequorea ma
15	1099	86.1	238	2	Q8WTC9	Q8wtc9 aequorea ma
16	1096	85.9	238	2	Q8WTC7	Q8wtc7 aequorea ma
17	1095	85.8	238	2	Q8WTC5	Q8wtc5 aequorea ma
18	636	49.8	234	2	Q6RYS7	Q6rys7 phialidium
19	477	37.4	225	2	Q6RYS5	Q6rys5 anthomedusa
20	436	34.2	262	2	Q6RYS6	Q6rys6 anthomedusa
21	252.5	19.8	226	2	Q8T6U0	Q8t6u0 dendronepht
22	251.5	19.7	225	2	Q95UA7	Q95ua7 montastraea
23	251.5	19.7	225	2	Q7Z0W5	Q7z0w5 montastraea
24	251	19.7	232	2	Q6RYS4	Q6rys4 anthomedusa
25	249	19.5	225	2	Q963F5	Q963f5 montastraea
26	247.5	19.4	225	2	Q7Z0W9	Q7z0w9 montastraea
27	247	19.4	230	2	Q66PW1	Q66pw1 scolymia cu
28	243.5	19.1	266	2	Q9U6Y3	Q9u6y3 clavularia
29	241.5	18.9	266	2	Q6USK3	Q6usk3 montastraea
30	239	18.7	225	2	Q8I6J8	Q8i6j8 trachyphyll
31	237	18.6	224	2	Q8MU48	Q8mu48 montastraea
32	237	18.6	225	2	Q66ND3	Q66nd3 montastraea
33	236.5	18.5	225	2	Q8T5F1	Q8t5f1 montastraea
34	235	18.4	225	2	Q7Z0W4	Q7z0w4 montastraea
35	222	17.4	231	2	Q66PV5	Q66pv5 acropora mi
36	221.5	17.4	221	2	Q66PV1	Q66pv1 acropora te
37	221.5	17.4	223	2	Q6R8F5	Q6r8f5 astrangia l
38	220	17.2	227	2	Q66ND4	Q66nd4 montastraea
39	219	17.2	219	2	Q66ND2	Q66nd2 montastraea
40	219	17.2	227	2	Q66ND5	Q66nd5 montastraea
41	218.5	17.1	221	2	Q66PU8	Q66pu8 acropora ac
42	218.5	17.1	259	2	Q8MMA2	Q8mma2 agaricia fr
43	218	17.1	239	2	Q8MMA1	Q8mma1 agaricia ag
44	216	16.9	227	2	Q7Z0W6	Q7z0w6 montastraea
45	216	16.9	231	2	Q66PV3	Q66pv3 acropora no

```
                           ALIGNMENTS

RESULT 1
Q8GHE2
ID   Q8GHE2       PRELIMINARY;        PRT;        238 AA.
AC   Q8GHE2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescence protein.
GN   Name=2289Gfp;
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM2289;
RA   Koranyi P., Berenyi M., Burg K.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF324408; AAN86140.1; -.
     HSSP; P42212; 1GFL.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD011756; Green_fl_protein; 1.
SQ   SEQUENCE  238 AA;  26914 MW;  F84840F1F9064018 CRC64;

Query Match             98.9%;  Score 1262;  DB 2;  Length 238;
Best Local Similarity   98.7%;  Pred. No. 1.1e-93;
Matches  235;  Conservative   1; Mismatches   2; Indels   0;  Gaps   0;

Qy        1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy       61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
            ||| :|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy      121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db      121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy      181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238

RESULT 2
GFP_AEQVI
ID   GFP_AEQVI    STANDARD;           PRT;        238 AA.
AC   P42212; Q17104; Q27903;
```

```
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Green fluorescent protein.
GN   Name=GFP;
OS   Aequorea victoria (Jellyfish).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=6100;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=9175527; PubMed=1347277; DOI=10.1016/0378-1119(92)90691-H;
RA   Prasher D.C., Eckenrode V.K., Ward W.W., Prendergast F.G.,
RA   Cormier M.J.;
RT   "Primary structure of the Aequorea victoria green-fluorescent
RT   protein.";
RL   Gene 111:229-233(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94185810; PubMed=8137953; DOI=10.1016/0014-5793(94)80472-9;
RA   Inouye S., Tsuji F.I.;
RT   "Aequorea green fluorescent protein. Expression of the gene and
RT   fluorescence characteristics of the recombinant protein.";
RL   FEBS Lett. 341:277-280(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97299832; PubMed=9154981; DOI=10.1023/A:1005740823703;
RA   Rouwendal G.J.A., Mendes O., Wolbert E.J.H., de Boer A.D.;
RT   "Enhanced expression in tobacco of the gene encoding green fluorescent
RT   protein by modification of its codon usage.";
RL   Plant Mol. Biol. 33:989-999(1997).
RN   [4]
RP   CHROMOPHORE.
RX   MEDLINE=9192221; PubMed=8448132;
RA   Cody C.W., Prasher D.C., Westler W.M., Prendergast F.G., Ward W.W.;
RT   "Chemical structure of the hexapeptide chromophore of the Aequorea
RT   green-fluorescent protein.";
RL   Biochemistry 32:1212-1218(1993).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=96355665; PubMed=8703075;
RA   Ormoe M., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y.,
RA   Remington S.J.;
RT   "Crystal structure of the Aequorea victoria green fluorescent
RT   protein.";
RL   Science 273:1392-1395(1996).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=98294543; PubMed=9631087;
RA   Yang F., Moss L.G., Phillips G.N. Jr.;
RT   "The molecular structure of green fluorescent protein.";
RL   Nat. Biotechnol. 14:1246-1251(1996).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF MUTANT WITH YELLOW EMISSION.
RX   MEDLINE=98455509; PubMed=9782051; DOI=10.1016/S0969-2126(98)00127-0;
RA   Wachter R.M., Elsliger M.A., Kallio K., Hanson G.T., Remington S.J.;
RT   "Structural basis of spectral shifts in the yellow-emission variants
RT   of green fluorescent protein.";
RL   Structure 6:1267-1277(1998).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX   MEDLINE=99238303; PubMed=10220315; DOI=10.1021/bi9902182;
RA   Elsliger M.A., Wachter R.M., Hanson G.T., Kallio K., Remington S.J.;
RT   "Structural and spectral response of green fluorescent protein
RT   variants to changes in pH.";
RL   Biochemistry 38:5296-5301(1999).
CC   -!- FUNCTION: Energy-transfer acceptor. Its role is to transduce the
CC       blue chemiluminescence of the protein aequorin into green
CC       fluorescent light by energy transfer. Fluoresces in vivo upon
CC       receiving energy from the Ca(2+)-activated photoprotein aequorin.
CC       Absorbs light maximally at 395 nm and exhibits a smaller
CC       absorbance peak at 470 nm. The fluorescence emission spectrum
CC       peaks at 509 nm with a shoulder at 540 nm.
CC   -!- SUBUNIT: Monomer.
CC   -!- TISSUE SPECIFICITY: Photocytes.
CC   -!- PTM: Contains a covalently attached chromophore, which is composed
CC       of modified amino acid residues. The chromophore is formed upon
CC       cyclization of the residues Ser-dehydroTyr-Gly.
CC   -!- BIOTECHNOLOGY: Has become a useful and ubiquitous tool for making
CC       chimeric proteins of GFP linked to other proteins where it
CC       functions as a fluorescent protein tag. GFP tolerates N-and C-
CC       terminal fusion to a broad variety of proteins. It has been
CC       expressed in bacteria, yeast, slime mold, plants, Drosophila,
CC       zebrafish, and in mammalian cells. As a noninvasive fluorescent
CC       marker in living cells, it allows for a wide range of applications
CC       where it may function as a cell lineage tracer, reporter of gene
CC       expression, or as a measure of protein-protein interactions.
CC   -!- DATABASE: NAME=Protein Spotlight; NOTE=Issue 11 of June 2001;
CC       WWW="http://www.expasy.org/spotlight/articles/sptlc011.html".
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M62654; AAA27722.1; -.
DR   EMBL; M62653; AAA27721.1; -.
DR   EMBL; L29345; AAA58846.1; -.
DR   EMBL; X96418; CAA65278.1; -.
DR   PIR; JS0692; JQ1514.
DR   PDB; 1B9C; X-ray; A/B/C/D=1-238.
DR   PDB; 1BFP; X-ray; @=1-238.
DR   PDB; 1C4F; X-ray; A=1-238.
DR   PDB; 1CV7; X-ray; @=-.
DR   PDB; 1EMA; X-ray; @=1-238.
DR   PDB; 1EMB; X-ray; A/B/C/D=1-238.
DR   PDB; 1EMC; X-ray; A/B/C/D=1-238.
DR   PDB; 1EME; X-ray; @=1-238.
DR   PDB; 1EMF; X-ray; @=1-238.
DR   PDB; 1EMG; X-ray; A=1-238.
DR   PDB; 1EMK; X-ray; @=1-238.
DR   PDB; 1EML; X-ray; A=1-238.
DR   PDB; 1EMM; X-ray; @=1-238.
DR   PDB; 1F09; X-ray; A=1-238.
DR   PDB; 1F0B; X-ray; A=1-238.
DR   PDB; 1GFL; X-ray; A/B=1-238.
DR   PDB; 1H6R; X-ray; A/B/C=1-238.
DR   PDB; 1HCJ; X-ray; A/B/C/D=1-238.
DR   PDB; 1HUY; X-ray; A=1-238.
DR   PDB; 1JBY; X-ray; A=1-238.
DR   PDB; 1JBZ; X-ray; A=1-238.
DR   PDB; 1JC0; X-ray; A/B/C=1-238.
DR   PDB; 1JC1; X-ray; A/B/C=1-238.
DR   PDB; 1KP5; X-ray; A/B=1-237.
DR   PDB; 1KYP; X-ray; A=1-238.
DR   PDB; 1KYR; X-ray; A=1-238.
DR   PDB; 1KYS; X-ray; A=1-238.
DR   PDB; 1MYW; X-ray; A=1-238.
DR   PDB; 1Q4A; X-ray; A=1-238.
DR   PDB; 1Q4B; X-ray; A=1-238.
DR   PDB; 1Q4C; X-ray; A=1-238.
DR   PDB; 1Q4D; X-ray; A=1-238.
DR   PDB; 1Q4E; X-ray; A=1-238.
DR   PDB; 1Q73; X-ray; A=1-238.
DR   PDB; 1QXT; X-ray; A=1-229.
DR   PDB; 1QY3; X-ray; A=1-229.
DR   PDB; 1QYF; X-ray; A=1-229.
DR   PDB; 1QYO; X-ray; A=1-238.
DR   PDB; 1QYQ; X-ray; A=1-238.
DR   PDB; 1YFP; X-ray; A/B=3-228.
DR   PDB; 2EMD; X-ray; @=1-238.
DR   PDB; 2EMN; X-ray; @=1-238.
DR   PDB; 2EMO; X-ray; @=1-238.
```

```
RESULT 3
Q71RY9
ID   Q71RY9              PRELIMINARY;      PRT;    238 AA.
AC   Q71RY9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Green fluorescence protein.
GN   Name=289Gfp;
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CCM289;
RA   Koranyi P., Berenyi M., Burg K.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF324407; AAN86139.1; -.
DR   HSSP; P42212; 1B9C.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
SQ   SEQUENCE 238 AA; 26886 MW; EA5A6F21FBFB6E05 CRC64;

Query Match             98.6%;  Score 1258;  DB 2;  Length 238;
Best Local Similarity   98.3%;  Pred. No. 2.3e-93;
Matches  234;  Conservative   2;  Mismatches   2;  Indels    0;  Gaps    0;

Qy    1 MSKGEBLFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEBLFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQRTIFFKDDGNYKTRAEVKFEGDTLV 120
          ||  :|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQRTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 4
Q8GHE4
ID   Q8GHE4              PRELIMINARY;      PRT;    238 AA.
AC   Q8GHE4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescence protein.
GN   Name=375Gfp;
OS   Azomonas agilis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azomonas.
OX   NCBI_TaxID=116849;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Koranyi P., Berenyi M., Burg K.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF324405; AAN86137.1; -.
DR   HSSP; P42212; 1BFP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
KW   3D-structure; Direct protein sequencing; Luminescence.
KW   5-imidazolinone (Ser-Gly).
FT   CROSSLNK     65     67
FT   MOD_RES      66     66      (Z)-2,3-didehydrotyrosine.
FT   VARIANT      66     66      F -> Y.
FT   VARIANT     100    100      T -> S.
FT   VARIANT     108    108      L -> M.
FT   VARIANT     141    141      V -> I.
FT   VARIANT     219    219
FT   CONFLICT     25     25      S -> G (in Ref. 3).
FT   CONFLICT     80     80      H -> Q (in Ref. 2).
FT   CONFLICT    157    157      Q -> R (in Ref. 3).
FT   CONFLICT    172    172      Q -> P (in Ref. 2).
FT   CONFLICT    172    172      E -> K (in Ref. 2).
FT   TURN          3      4
FT   HELIX        11     23
FT   STRAND       25     37
FT   TURN         40     41
FT   HELIX        49     57
FT   TURN         61     69
FT   STRAND       73     76
FT   HELIX        83     87
FT   TURN         92    101
FT   STRAND      105    116
FT   TURN        118    132
FT   STRAND      135    141
FT   HELIX       148    156
FT   TURN        159    160
FT   STRAND      172    173
FT   TURN        176    199
FT   STRAND      208    211
FT   TURN        217    227
SQ   SEQUENCE 238 AA; 26886 MW; EA5A6F21FBFB6E05 CRC64;

Query Match             98.6%;  Score 1258;  DB 1;  Length 238;
Best Local Similarity   98.3%;  Pred. No. 2.3e-93;
Matches  234;  Conservative   2;  Mismatches   2;  Indels    0;  Gaps    0;

Qy    1 MSKGEBLFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSKGEBLFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQRTIFFKDDGNYKTRAEVKFEGDTLV 120
          ||| :|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQRTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
```

```
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
SQ   SEQUENCE   238 AA;  26902 MW;  15FE9B9C5B4F6B89 CRC64;

Query Match             98.5%;  Score 1257;  DB 2;  Length 238;
Best Local Similarity   98.3%;  Pred. No. 2.8e-93;
Matches  234;  Conservative   1;  Mismatches   3;  Indels   0;  Gaps   0;

Qy       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy      61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy     121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NRIELKGTDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 5
Q8GHE3
ID   Q8GHE3            PRELIMINARY;        PRT;    238 AA.
AC   Q8GHE3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescence protein.
GN   Name=85Gfp;
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Koranyi P., Berenyi M., Burg K.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF324406; AAN86138.1; -.
DR   HSSP; P42212; 1GFL.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
SQ   SEQUENCE   238 AA;  26887 MW;  E0E1616BD2AF6188 CRC64;

Query Match             98.4%;  Score 1256;  DB 2;  Length 238;
Best Local Similarity   98.3%;  Pred. No. 3.4e-93;
Matches  234;  Conservative   1;  Mismatches   3;  Indels   0;  Gaps   0;

Qy       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLTFICTTGKLPVPWPTL 60

Qy      61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy     121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 6
Q93125
ID   Q93125            PRELIMINARY;        PRT;    238 AA.
AC   Q93125;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescence protein mutant 3.
GN   Name=GFP;
OS   Aequorea victoria (Jellyfish).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=6100;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96305137; PubMed=8707053; DOI=10.1016/0378-1119(95)00685-0;
RA   Cormack B.P., Valdivia R.H., Falkow S.;
RT   "FACS-optimized mutants of the green fluorescent protein (GFP).";
RL   Gene 173:33-38(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97195776; PubMed=9043107;
RA   Cormack B.P., Bertram G., Egerton M., Gow N.A.R., Falkow S.,
RA   Brown A.J.P.;
RT   "Yeast-enhanced green fluorescent protein (yEGFP) a reporter of gene
RT   expression in Candida albicans.";
RL   Microbiology 143:303-311(1997).
DR   EMBL; U73901; AAB18957.1; -.
DR   PDB; 2YFP; X-ray; A=1-238.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
SQ   SEQUENCE   238 AA;  26840 MW;  A28622809A9DEA60 CRC64;

Query Match             98.1%;  Score 1252;  DB 2;  Length 238;
Best Local Similarity   97.9%;  Pred. No. 7.1e-93;
Matches  233;  Conservative   2;  Mismatches   3;  Indels   0;  Gaps   0;

Qy       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy      61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           ||||:||||||||:|:|||||||||||||||||||||||||||||||||||||||||||
Db      61 VTTFGYGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy     121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180

Qy     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238


RESULT 7
Q17105
ID   Q17105            PRELIMINARY;        PRT;    238 AA.
AC   Q17105;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescent protein (Fragment).
GN   Name=GFP;
OS   Aequorea victoria (Jellyfish).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=6100;
```

RN   [1]
RP   SEQUENCE FROM N.A.
RA   Watkins J.N., Campbell A.K.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X83959; CAA58789.1; -.
DR   PIR; JS0692; JQ1514.
DR   HSSP; P42212; 1GFL.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
FT   NON_TER   238   238
SQ   SEQUENCE   238 AA; 26950 MW; 26E2B8450E748E44 CRC64;

Query Match   95.3%; Score 1216; DB 2; Length 238;
Best Local Similarity   94.1%; Pred. No. 5.6e-90; Indels 0; Gaps 0;
Matches 224; Conservative 7; Mismatches 7;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ::|||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPVLVELDGDVNGQKFSVSGEGEGDATYGKLTLNFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||| ::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFYKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              |||||||||||||||||||||:||||||||||||:|:|||||||||||||||::||||||
Db   121 NRIELKGIDFKEDGNILGHKMEYNYNSHNVYIMGDKPKNGIKVNFKIRHNIKDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||||:|:||||||||:||||||||:||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSQDPHGKRDHMVLLEFVTSAGITHGMDELYK 238

RESULT 8
Q17106
ID   Q17106   PRELIMINARY;   PRT;   238 AA.
AC   Q17106;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescent protein (Fragment).
GN   Name=GFP;
OS   Aequorea victoria (Jellyfish).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=6100;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Watkins J.N., Campbell A.K.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X83960; CAA58790.1; -.
DR   PIR; JS0692; JQ1514.
DR   HSSP; P42212; 1BFP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
FT   NON_TER   238   238
SQ   SEQUENCE   238 AA; 26867 MW; BD4648262D8BABD4 CRC64;

Query Match   94.1%; Score 1201; DB 2; Length 238;
Best Local Similarity   93.3%; Pred. No. 9e-89; Indels 0; Gaps 0;
Matches 222; Conservative 8; Mismatches 8;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||||||||||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db   1 MSKGEELFTGVVPILVELDGDVNGQKFSVRGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||| ::||||||||||||||||||:|||||||||||||:||||||||||||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKQHDFLKSAMPEGYVQERTIFYKDDGNYKTRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              ||||||||||||||||||||:||||||||||||:|:|||||||||||||||::||||||
Db   121 NRIELKGIDFKEDGNILGHKMEYNYNSHNVYIMGDKPKNGIKVNFKIRHNIKDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
              ||||||||||||||||||||||||||:|:||||||||:|||||||:||||||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSQDPHGKRDHMVLLEFVTSAGITHGMDELYK 238

RESULT 9
Q6YGZ0
ID   Q6YGZ0   PRELIMINARY;   PRT;   238 AA.
AC   Q6YGZ0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Green fluorescent protein.
OS   Aequorea coerulescens (belt jellyfish).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=210840;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22726112; PubMed=12693991; DOI=10.1042/BJ20021966;
RA   Gurskaya N.G., Fradkov A.F., Pounkova N.I., Staroverov D.B.,
RA   Bulina M.E., Yanushevich Y.G., Labas Y.A., Lukyanov S., Lukyanov K.A.;
RT   "A colourless green fluorescent protein homologue from the non-
RT   fluorescent hydromedusa Aequorea coerulescens and its fluorescent
RT   mutants.";
RL   Biochem. J. 373:403-408(2003).
DR   EMBL; AY151052; AAN41637.1; -.
DR   HSSP; P42212; 1B9C.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   InterPro; IPR009017; GFP_like.
DR   InterPro; IPR011584; GFP_related.
DR   InterPro; IPR000786; Green_fl_protein.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
SQ   SEQUENCE   238 AA; 26896 MW; DE72EDB7ED9F9FE CRC64;

Query Match   93.1%; Score 1188; DB 2; Length 238;
Best Local Similarity   91.2%; Pred. No. 1e-87; Indels 0; Gaps 0;
Matches 217; Conservative 10; Mismatches 11;

Qy   1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
              ||||.|||||||||||:||.|||||||||||||||||||||||||||||||||||||||||
Db   1 MSKGAELFTGVVPILIELNGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60

Qy   61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
              ||| :|||||||||||||:||||||||||||||||||||||||||||:|||||||||||
Db   61 VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKSRAEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
              |||||.| :||||||||| |:|||||:||||||:||.|||||||||||||||||:||||||
Db   121 NRIELTGTDFKEDGNILGNKMEYNYNAHNVYIMTDKAKNGIKVNFKIRHNIEDGSVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
              |||||||||||||||||||||||||:|||||||||||||:.|.||||||:||||||||
Db   181 HYQQNTPIGDGPVLLPDNHYLSTQSTLSKDPNEKRDHMIYFBFVTAAAITHGMDELYK 238

RESULT 10
Q8WTC6
ID   Q8WTC6   PRELIMINARY;   PRT;   238 AA.
AC   Q8WTC6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

RESULT 11
Q8WP95
```
ID   Q8WP95
AC   Q8WP95;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Green fluorescent protein.
GN   Name=GFPxm;
OS   Aequorea macrodactyla.
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=147615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Qin Y.X., Pang S.Q.,
RA   Li S.J., Xia N.S.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY013824; AAK02062.1; -.
DR   EMBL; AY013821; AAK02059.1; -.
DR   HSSP; P42212; 1BFP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
SQ   SEQUENCE 238 AA; 27049 MW; 818SD0E5E529012B CRC64;

Query Match           86.6%;  Score 1105;  DB 2;  Length 238;
Best Local Similarity 83.2%;  Pred. No. 4.9e-81;
Matches  198;  Conservative  20;  Mismatches  20;  Indels   0;  Gaps   0;

Qy     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           ||||||||||:||||:|||||||:||||||:|||||||:||||:||||||||||||||||
Db     1 MSKGEELFTGIVPILIELDGDVHGHKFSVRGEGEGDADYGKLEIKFICTTGKLPVPWPTL 60

Qy    61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           ||||.||:||:||||:|||:.:|||||||||||:||||||:|||:||||:|||||||||||
Db    61 VTTLSYGIQCFARYPEHMKMNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           |||||||:||||||||||||||||:|:||||||.||:|||:||||||||||||:||||||
Db   121 NRIELKGMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
           |||.|.|:||||||:|.||||||||:|:||.||.|||||||||:.:.|||||||||||
Db   181 HYQTNVPLGDGPVLIPINHYLSTQTAISKDRNETRDHMVLEFFSACGHTHGMDELYK 238
```

RESULT 12
Q8WTC4
```
ID   Q8WTC4
AC   Q8WTC4;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescent protein.
GN   Name=GFP;
OS   Aequorea macrodactyla.
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=147615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA   Li S.J., Xia N.S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF435433; AAL33918.1; -.
DR   HSSP; P42212; 1KYP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
SQ   SEQUENCE 238 AA; 26956 MW; 7552lEAF0CBBA73A CRC64;

Query Match           86.4%;  Score 1102;  DB 2;  Length 238;
Best Local Similarity 84.5%;  Pred. No. 8.5e-81;
Matches  201;  Conservative  16;  Mismatches  21;  Indels   0;  Gaps   0;

Qy     1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
           ||||||||||:|||||||||||:||||||:|||||||:||||:||||||||||||||||
Db     1 MSKGEELFTGIVPILVELDGDVHGHKFSVRGEGEGDADYGKLEIKFICTTGKLPVPWPTL 60

Qy    61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
           ||||.||||:||:||||:||:.:|||||||||||:||||||:|||:||||:||||||||||
Db    61 VTTLGYGILCFARYPEHMKMNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120

Qy   121 NRIELKGIDFKEDGNILGHKLEYNYSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
           |||||||:||||||||||||||||:|:||||||.||:|||:||||||||||||:||||||
Db   121 NRIELKGMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQLAD 180

Qy   181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
           |||.|.|:||||||:|.||||||||:|:||.||.|||||||||:.:.|||||||||||
Db   181 HYQTNVPLGDGPVLIPINHYLSTQTAISKDRNETRDHMVLEFFSACGHTHGMDELYK 238
```

RESULT 13
Q8WTD0
```
ID   Q8WTD0
AC   Q8WTD0;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescent protein.
GN   Name=GFP;
OS   Aequorea macrodactyla.
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=147615;
```

RESULT 14

```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA   Li S.J., Xia N.S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF435427, AAL33912.1; -.
DR   HSSP; P42212; 1KYP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
DR   ProDom; PD013756; Green_fl_protein; 1.
SQ   SEQUENCE   238 AA;   26997 MW;   5F80A192173CB84D CRC64;

Query Match             86.2%;  Score 1100;  DB 2;  Length 238;
Best Local Similarity   83.2%;  Pred. No. 1.2e-80;
Matches  198;  Conservative  19;  Mismatches  21;  Indels  0;  Gaps  0;

Qy        1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
            ||||||||||::|:||||:|||||||||||||||||| |||||:|||||||||||||||||
Db        1 MSKGEELFTGIVPVLIELDGDVHGHKFSVRGEGEGDADYGKLEIKFICTTGKLPVPWPTL 60

Qy       61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRABVKFEGDTLV 120
            ||||:|||||:||:|||||  ||||||||||| |||||||| |||||||  |||||||||
Db       61 VTTLGYGIQCFARYPEHMKMNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120

Qy      121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
            |||||||||||||||||||||||| ||||||||  ||||||||||||||||| |||||||
Db      121 NRIELKGMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQLAD 180

Qy      181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
            ||| |||||||||| |:|||||:|:|||| ||| ||||:|::|| :|||||||||||||
Db      181 HYQTNVPLGDGPVLIPINHYLSLQTAISKDRNETRDHMVFLBFPSACGHTHGMDELYK 238
```

RESULT 15

```
ID   Q8WTC8               PRELIMINARY;        PRT;   238 AA.
AC   Q8WTC8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Green fluorescent protein.
GN   Name=GFP;
OS   Aequorea macrodactyla.
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC   Aequoreidae; Aequorea.
OX   NCBI_TaxID=147615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA   Li S.J., Xia N.S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF435429, AAL33914.1; -.
DR   HSSP; P42212; 1KYP.
DR   GO; GO:0006091; P:energy pathways; IEA.
DR   Pfam; PF01353; GFP; 1.
DR   PRINTS; PR01229; GFLUORESCENT.
SQ   SEQUENCE   238 AA;   27047 MW;   5F80A18FA1B7C84D CRC64;

Query Match             86.1%;  Score 1099;  DB 2;  Length 238;
Best Local Similarity   83.2%;  Pred. No. 1.5e-80;
Matches  198;  Conservative  19;  Mismatches  21;  Indels  0;  Gaps  0;

Qy        1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL 60
            ||||||||||::|:||||:|||||||||||||||||| |||||:|||||||||||||||||
Db        1 MSKGEELFTGIVPVLIELDGDVHGHKFSVRGEGEGDADYGKLEIKFICTTGKLPVPWPTL 60

Qy       61 VTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRABVKFEGDTLV 120
            ||||:|||||:||:|||||  ||||||||||| |||||||| |||||||  |||||||||
Db       61 VTTLGYGIQCFARYPEHMKMNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120

Qy      121 NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGGVQLAD 180
            ||||||||||||||||||||||||| |||||||| |||||||||||||||||| |||||||
Db      121 NRIELKGMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQLAD 180

Qy      181 HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
            ||| |||||||||| |:|||||:|:|||| ||| ||||:|::|| :|||||||||||||
Db      181 HYQTNVPLGDGPVLIPINHYLSFQTAISKDRNETRDHMVFLBFPSACGHTHGMDELYK 238
```

Search completed: September 16, 2005, 17:20:07
Job time : 115 secs